Computational social science

Social networks: dynamics

Course structure

Period IV

Week	Lecture	Exer. dl	Ext. dl	Topic
1	~~Feb 27~~	~~Mar 3~~	~~Mar 15~~	~~Introduction to CSS~~
2	~~Mar 6~~	~~Mar 10~~	~~Mar 22~~	~~Artificial societies & agent-based models~~
3	~~Mar 13~~	~~Mar 17~~	~~Mar 29~~	~~Data & digital traces~~
4	~~Mar 20~~	~~Mar 24~~	~~Apr 5~~	~~Counting things & analysing text~~
5	~~Mar 27~~	~~Mar 31~~	~~Apr 12~~	~~Social networks: structure~~
6	~~Apr 3~~	*	-	~~Introduction to the project~~

***Project deadline: May 26**
Project peer review: June 2

Period V

Week	Lecture	Exercise dl	Ext. dl	Topic
7	Apr 24	May 5	~~May 10~~	Ethics, privacy, legal
-	-	-	-	~~WAPPU~~
8	~~May 8~~	~~May 12**~~	May 24	~~Agent-based models & emergence~~
9	May 15	May 19***	May 31	Social networks: dynamics
10	May 22	May 26***	June 7	Experiments & interventions at scale
11	May 29	-	-	Computing for social good

****Bonus round**
*****Only lecture questions**

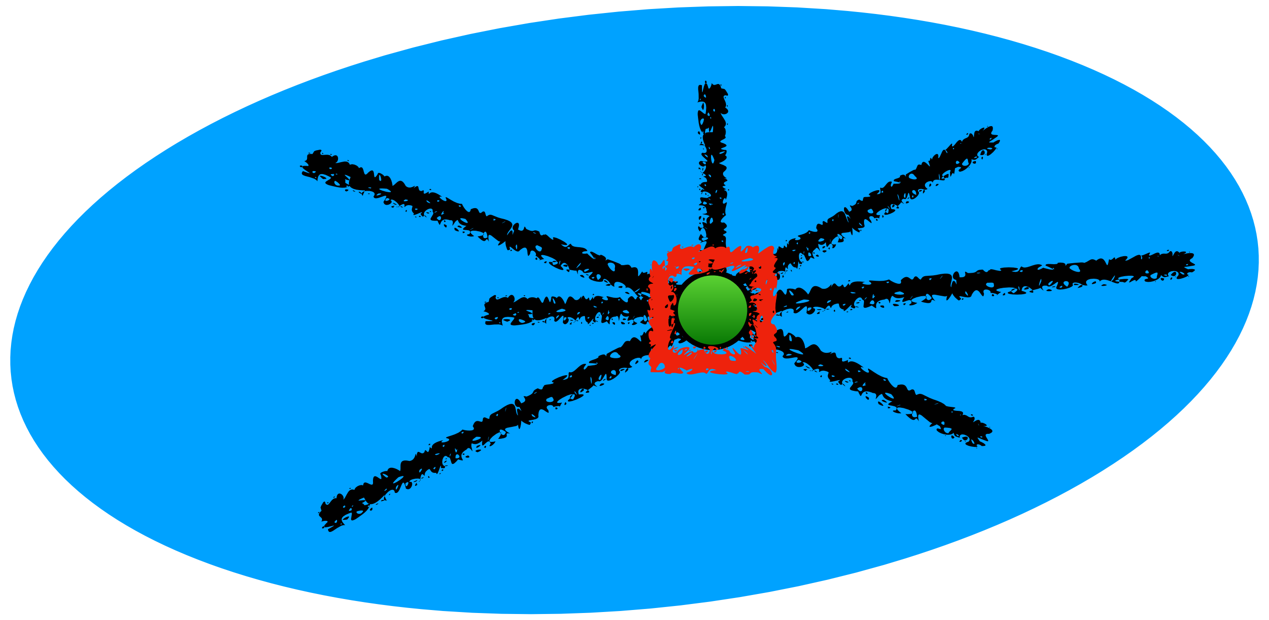
What kind of structure is there in social networks?

Why is the structure important?

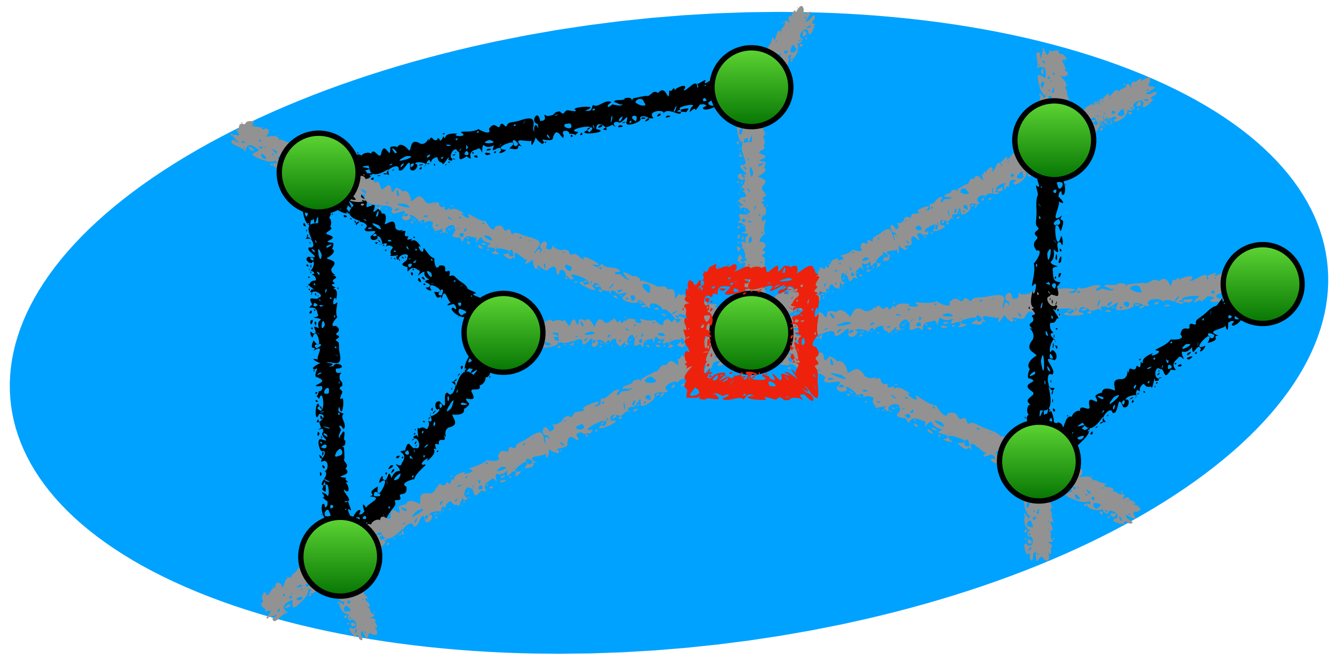
What can agent-based modelling of networks say about these questions?

Reminder: network structure

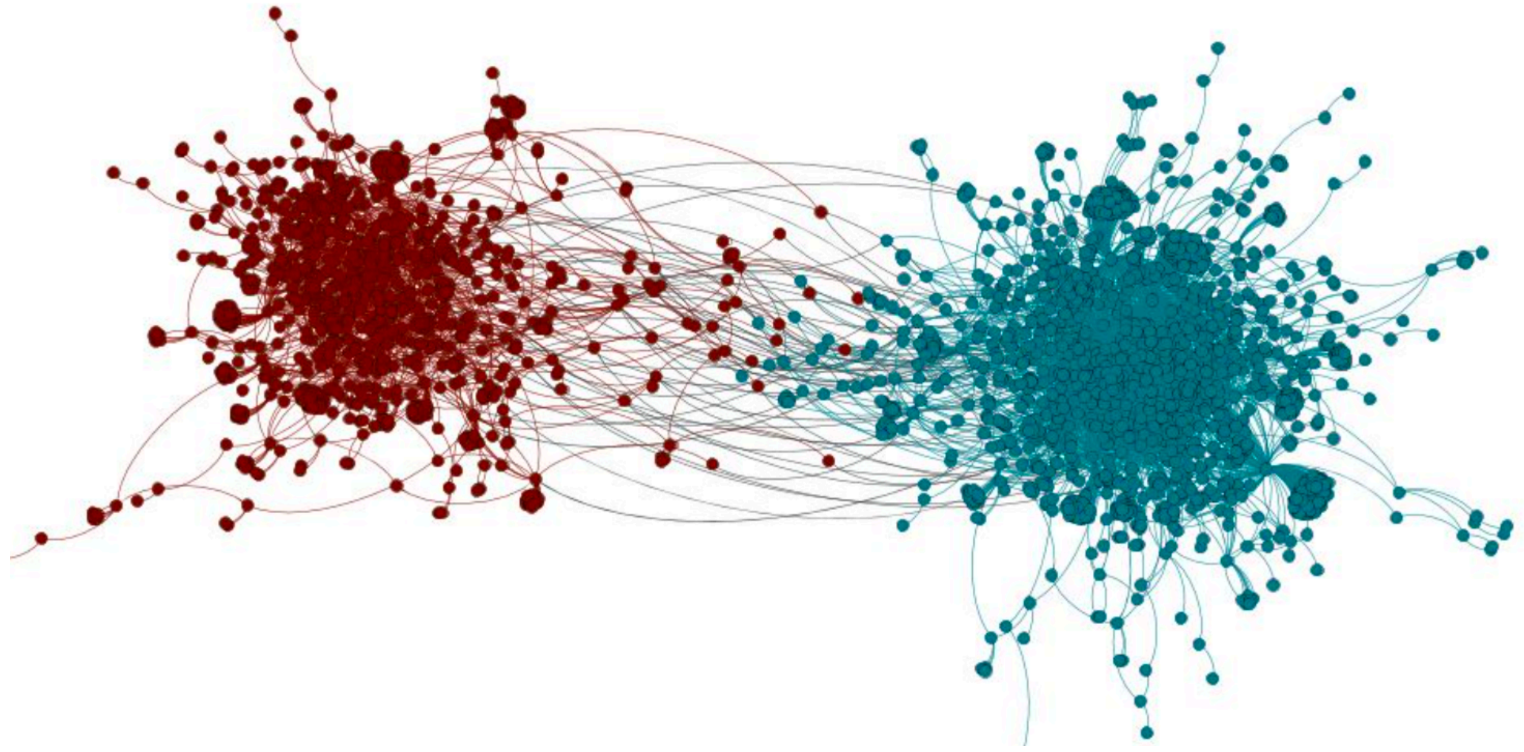
Degree heterogeneity



High clustering coefficients



Communities



What explains the emergence of the structure?

What explains the emergence of the structure?

How does the structure affect dynamics?

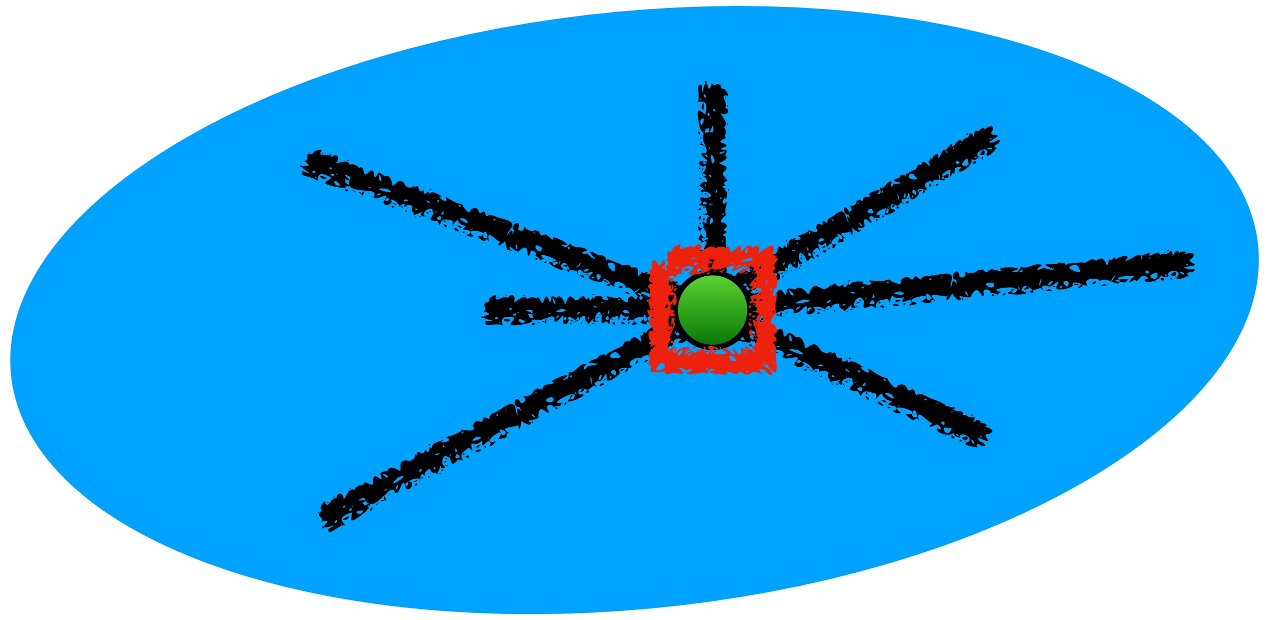
How are networks born?

They grow and evolve

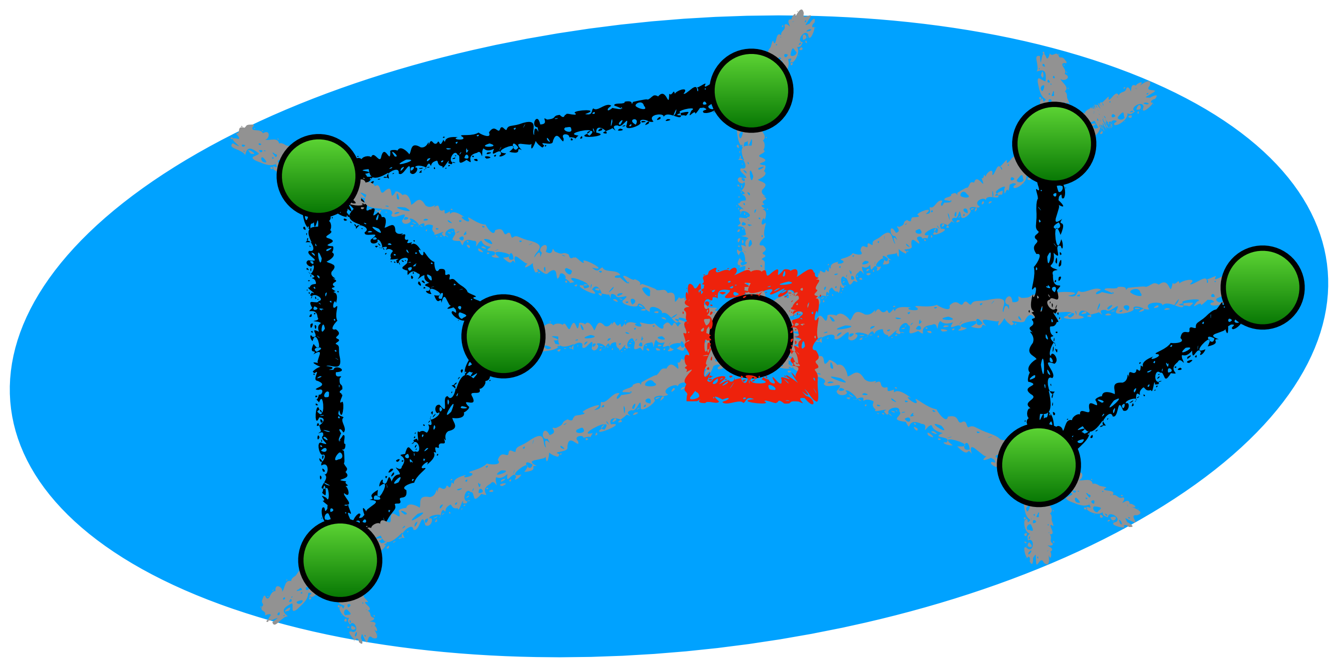
Dynamics → network structure

Network dynamics

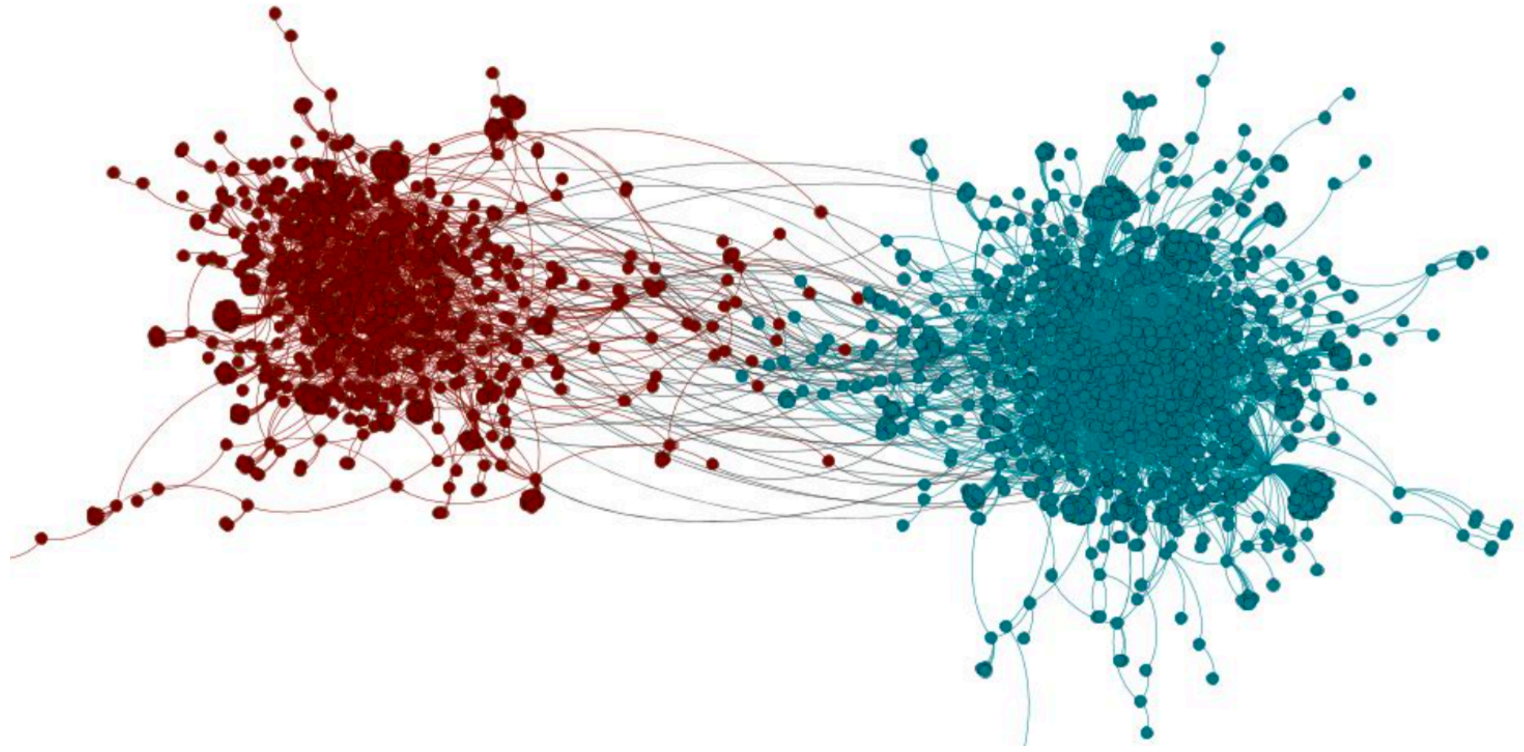
Degree heterogeneity



High clustering coefficients



Communities

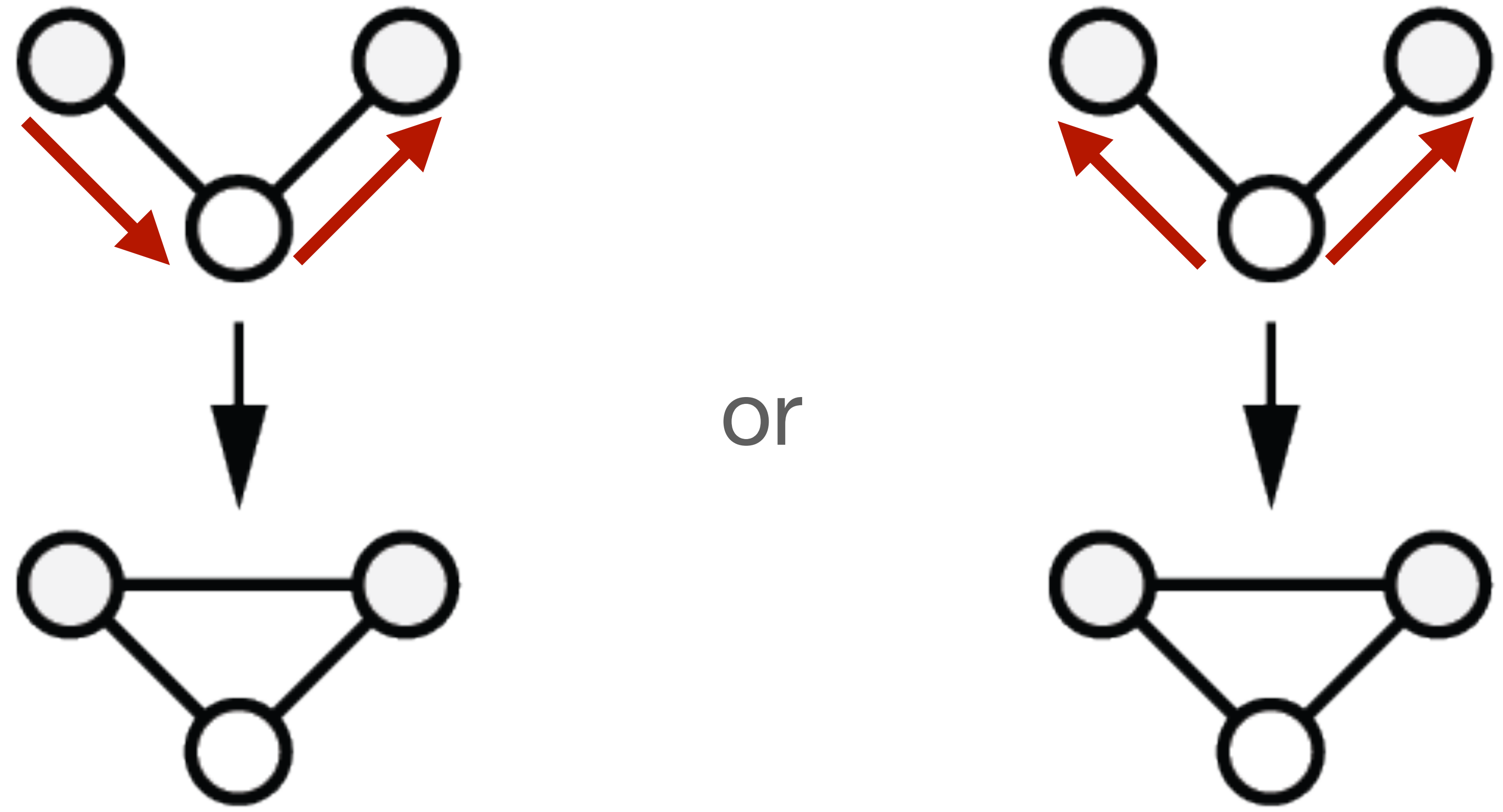


How do networks evolve?

Using the network itself to create new links

Creating new links

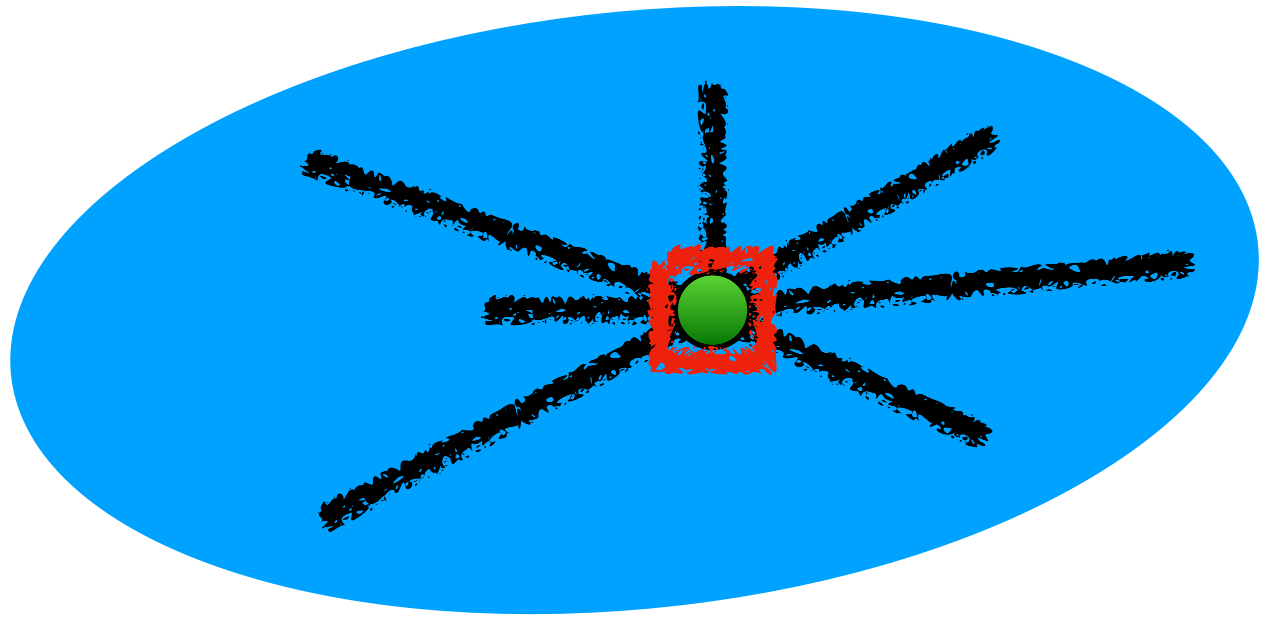
- *Social networks*:
  - You meet people based on your previous connections → **triadic closure**
    - Empirical evidence (e.g. Kossinets 2006)
    - Network models (e.g., Toivonen 2008)
    - Recommendation systems
- *Linked documents* (scientific articles, wikipedia pages, etc):
  - Find document based on the network and link to them
  - Create (partial) copies of old documents (or links in them)



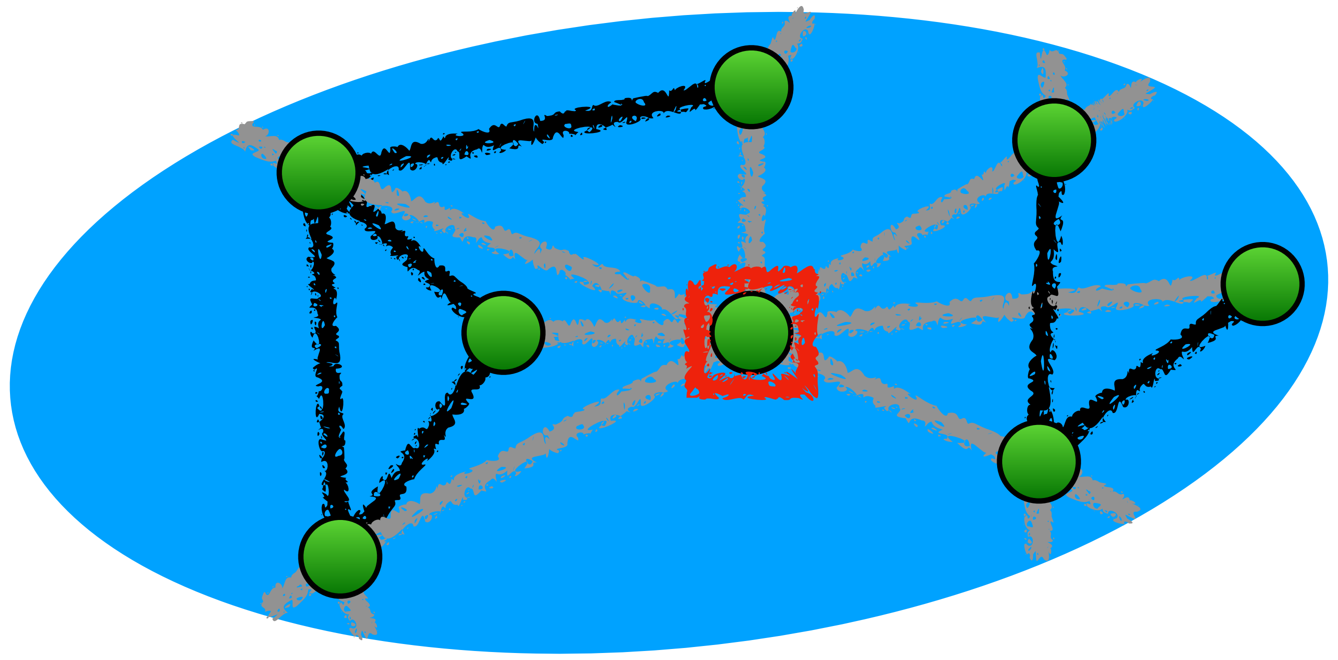
Dynamics → network structure

Network dynamics

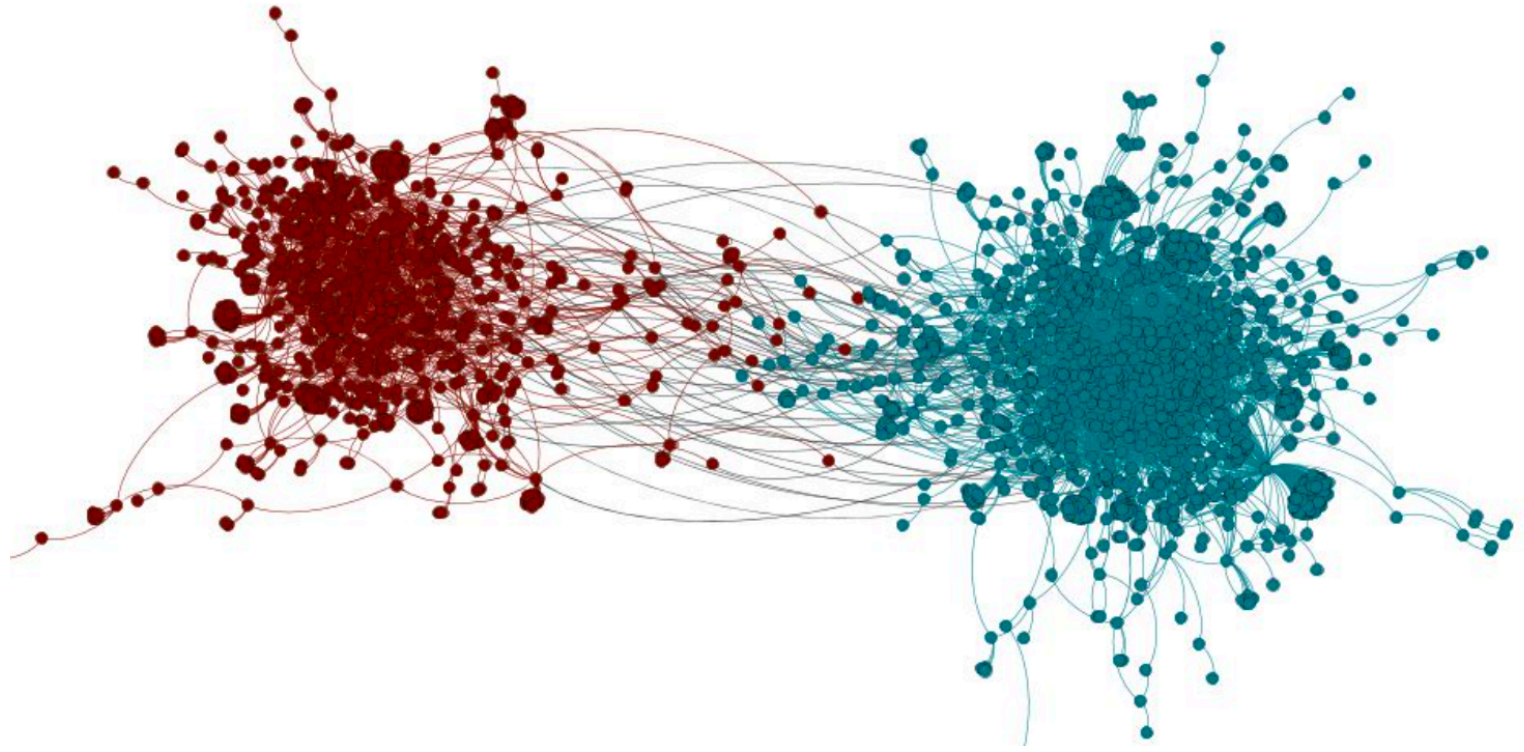
Degree heterogeneity



High clustering coefficients

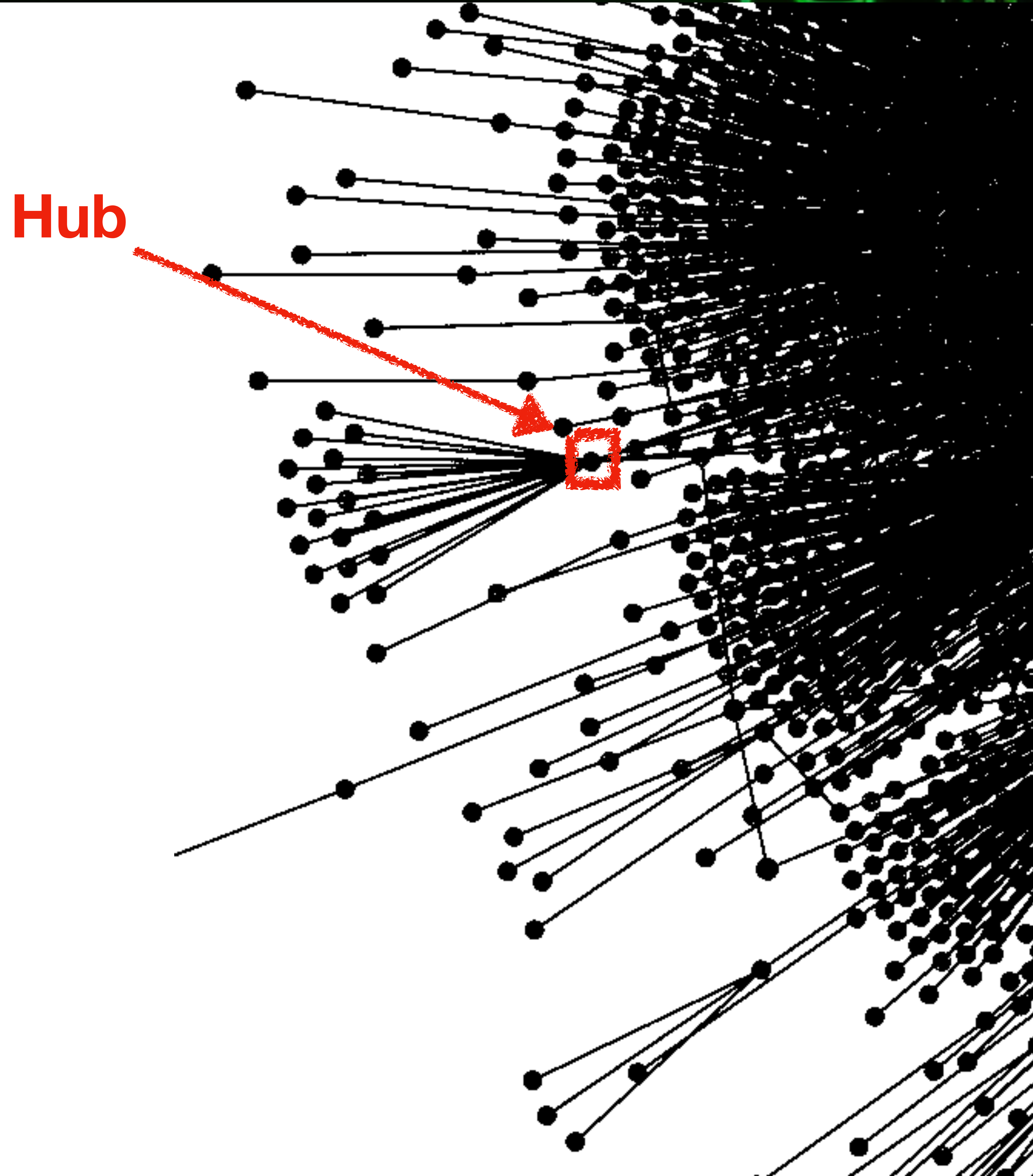


Communities



Reminder: degrees & hubs

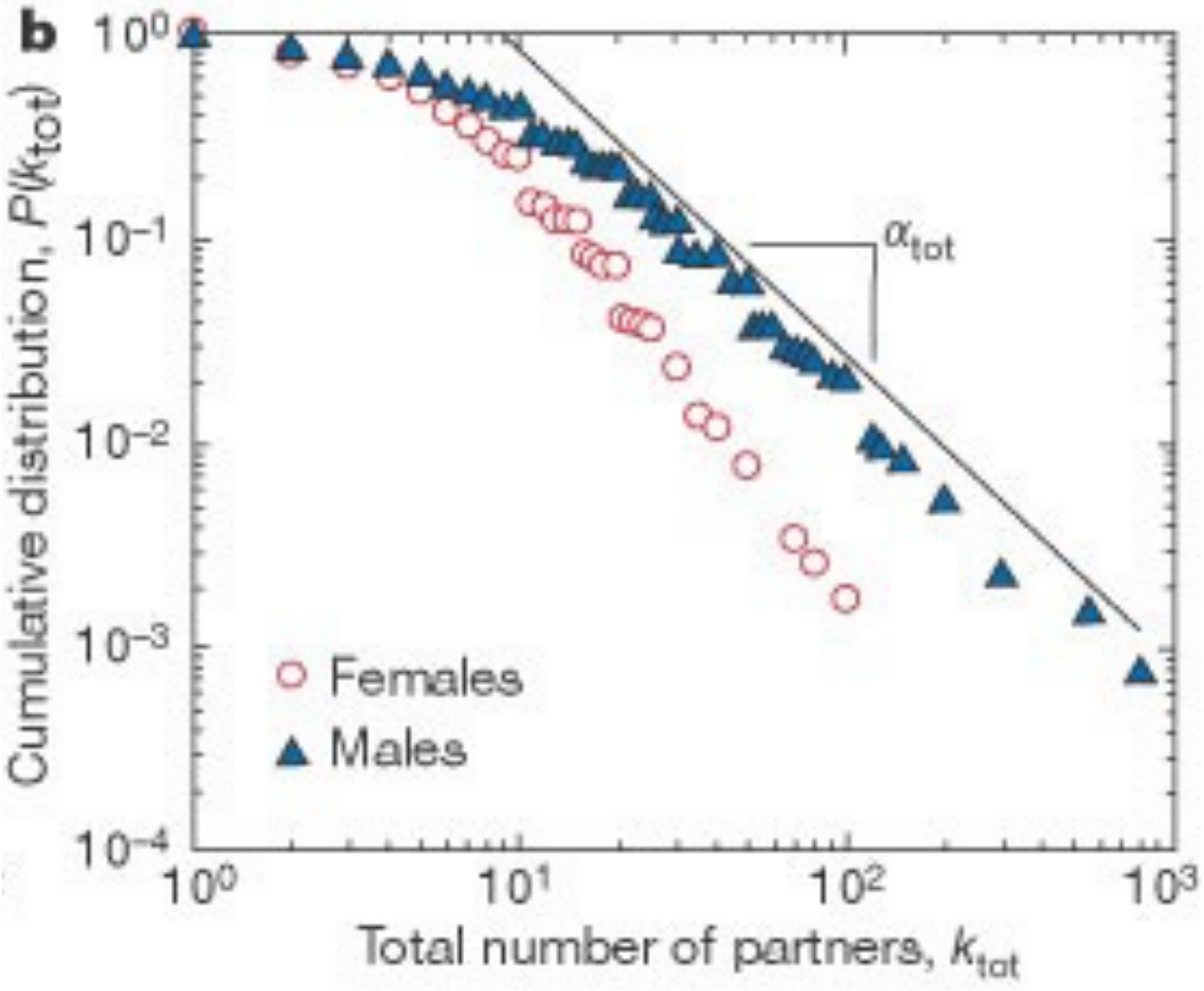
- Degree = k_i = number of connections/neighbors node i has
- Hub: a node that has very high degree
- Typically most nodes have very low degree, few hubs



Reminder: scale-free networks?

- Scale-free networks: power-law degree distribution: $P(k) \sim k^{-\alpha}$
- Often plotted in double logarithmic axis → straight line: $\log P(k) \sim -\alpha \log k$
- Many networks reported as scale-free in the literature

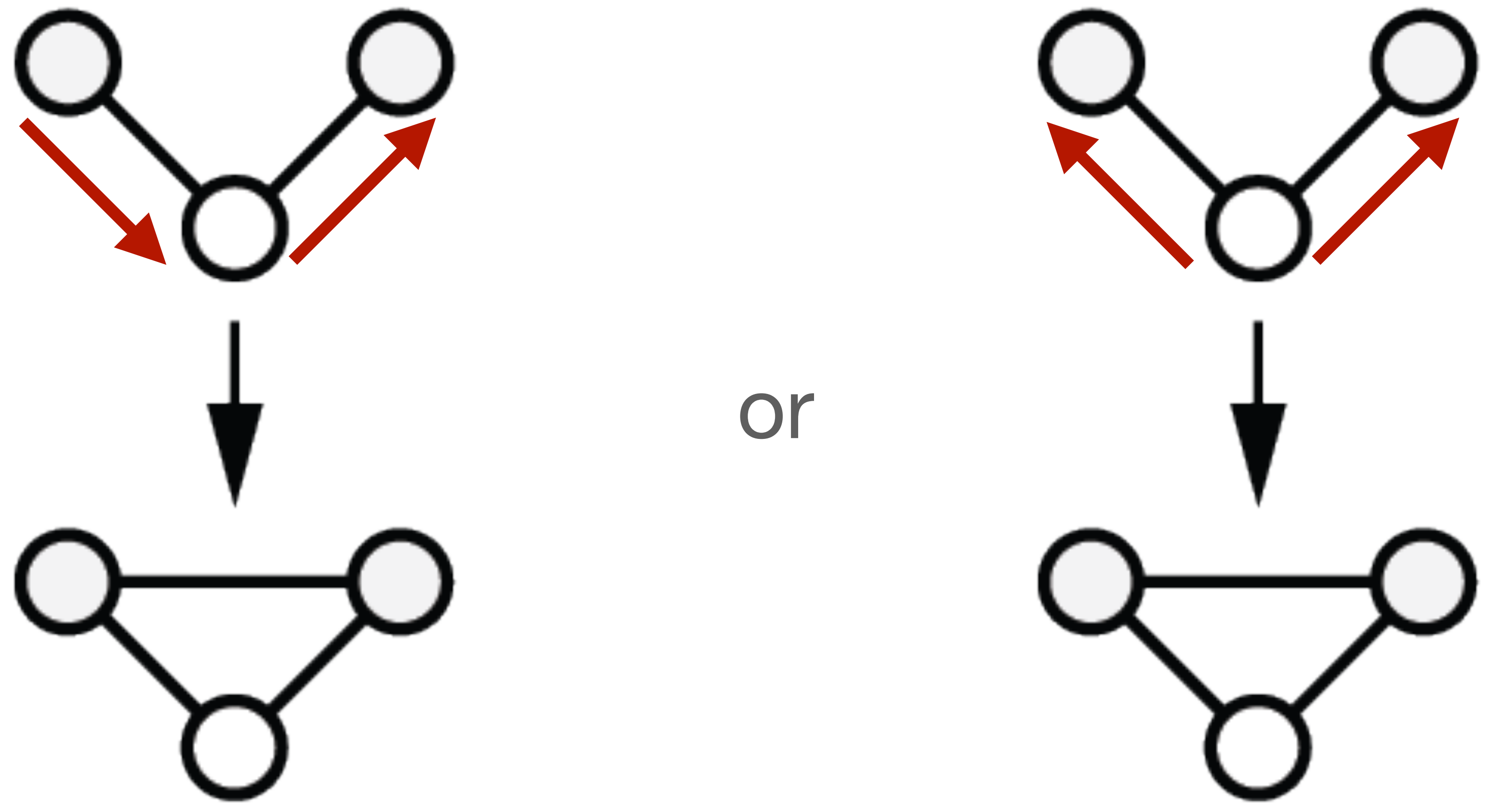
→ Typically in reality a "scale-free tail" (no power-law for low k values)



F. Liljeros et al., Nature 2001

Creating new links

- *Social networks*:
  - You meet people based on your previous connections → **triadic closure**
    - Empirical evidence (e.g. Kossinets 2006)
    - Network models (e.g., Toivonen 2008)
- *Linked documents* (scientific articles, wikipedia pages, etc):
  - Find document based on the network and link to them
  - Create (partial) copies of old documents (or links in them)

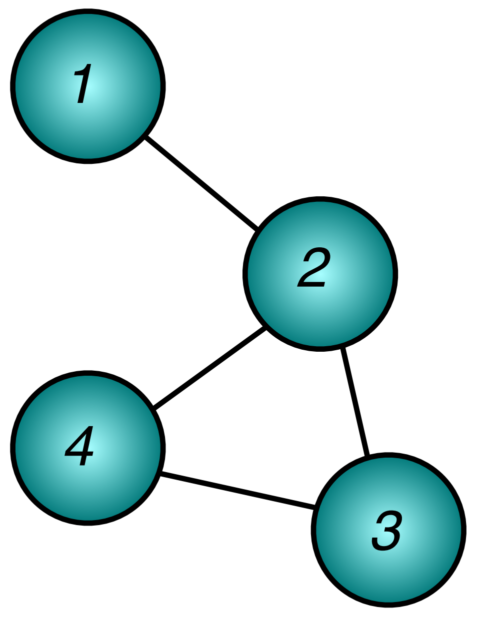


In all of these cases: Follow links to find nodes to link to

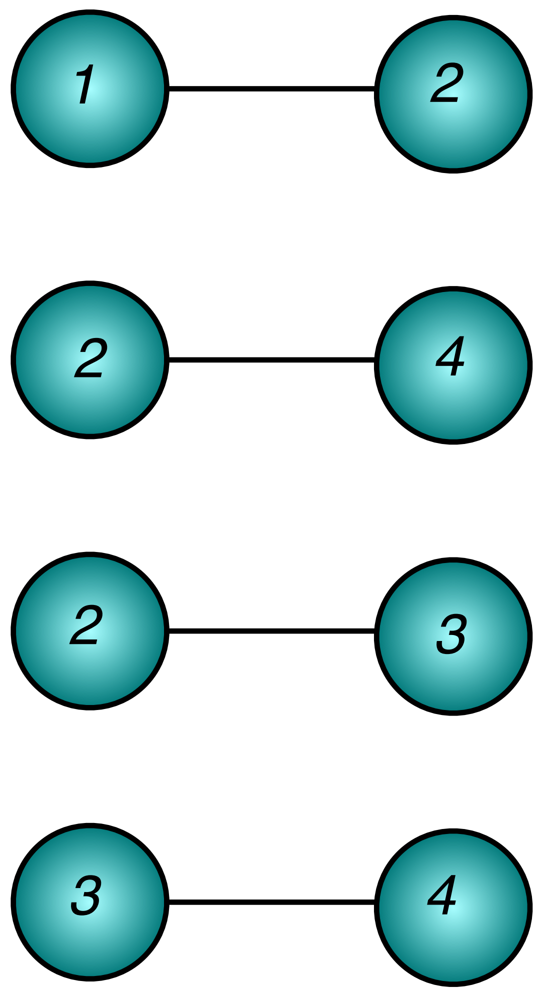
Reminder: Friendship paradox

If you follow a random link, what is the probability that node i with degree k_i is picked?

Example network:



Each link picked with probability 1/4



After that, each node is picked with probability 1/2

$$p_1 = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$$

$$p_2 = 3 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{8}$$

$$p_3 = 2 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{2}{8}$$

$$p_4 = 2 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{2}{8}$$

$$p(\text{'follow random link, reach node i'}) = \frac{k_i}{\sum_j k_j}$$

Following links

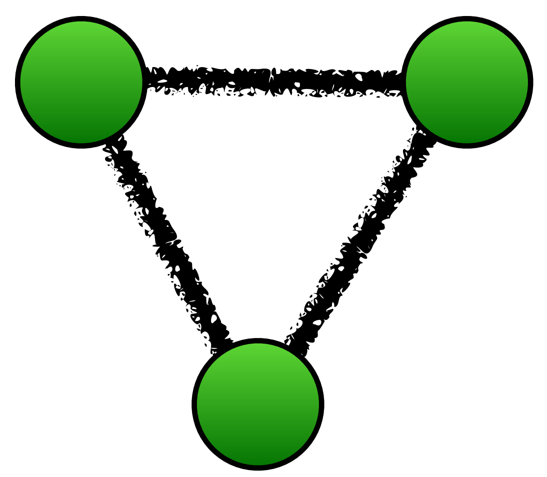
=

Preferentially choosing high degree nodes

Preferential attachment

- In preferential attachment models new links connect preferentially to high degree nodes
- Example: Barabási-Albert model:

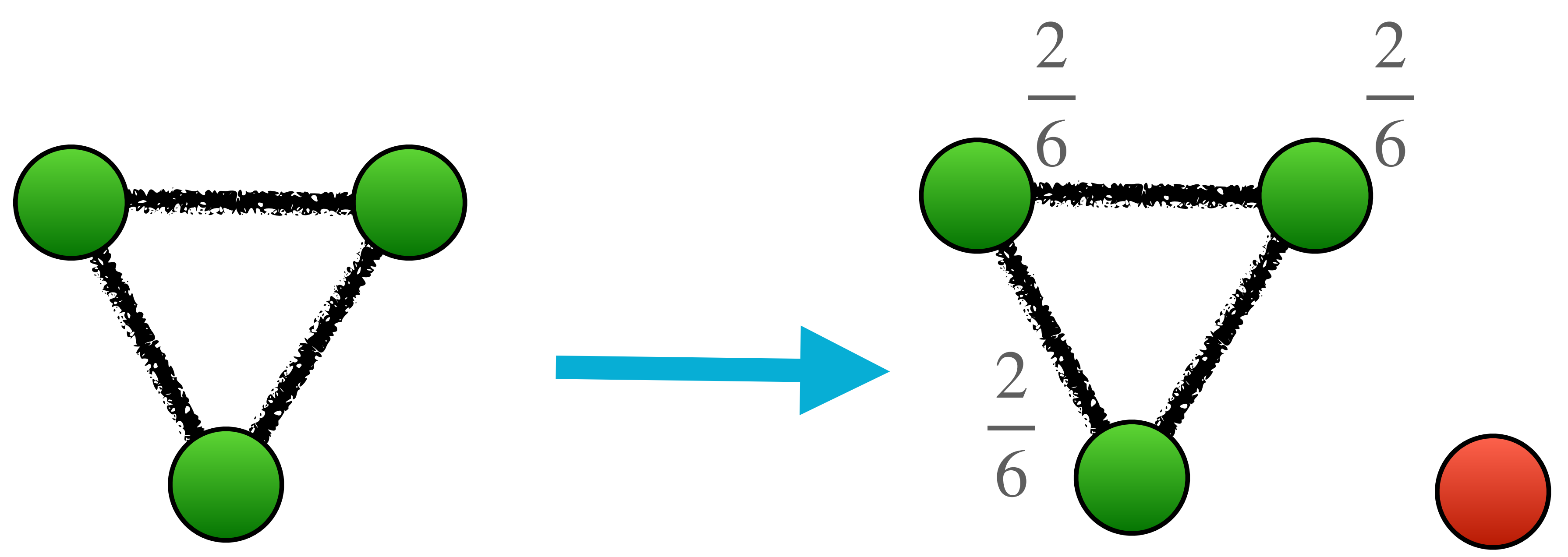
1. Start off with a small seed network
2. Choose *m* existing nodes with a probability directly proportional to their degree: $\frac{k_i}{\sum_j k_j}$
3. Connect a new node to the m chosen nodes
4. If not yeah *N* nodes, go to 2.



Preferential attachment

- In preferential attachment models new links connect preferentially to high degree nodes
- Example: Barabási-Albert model:

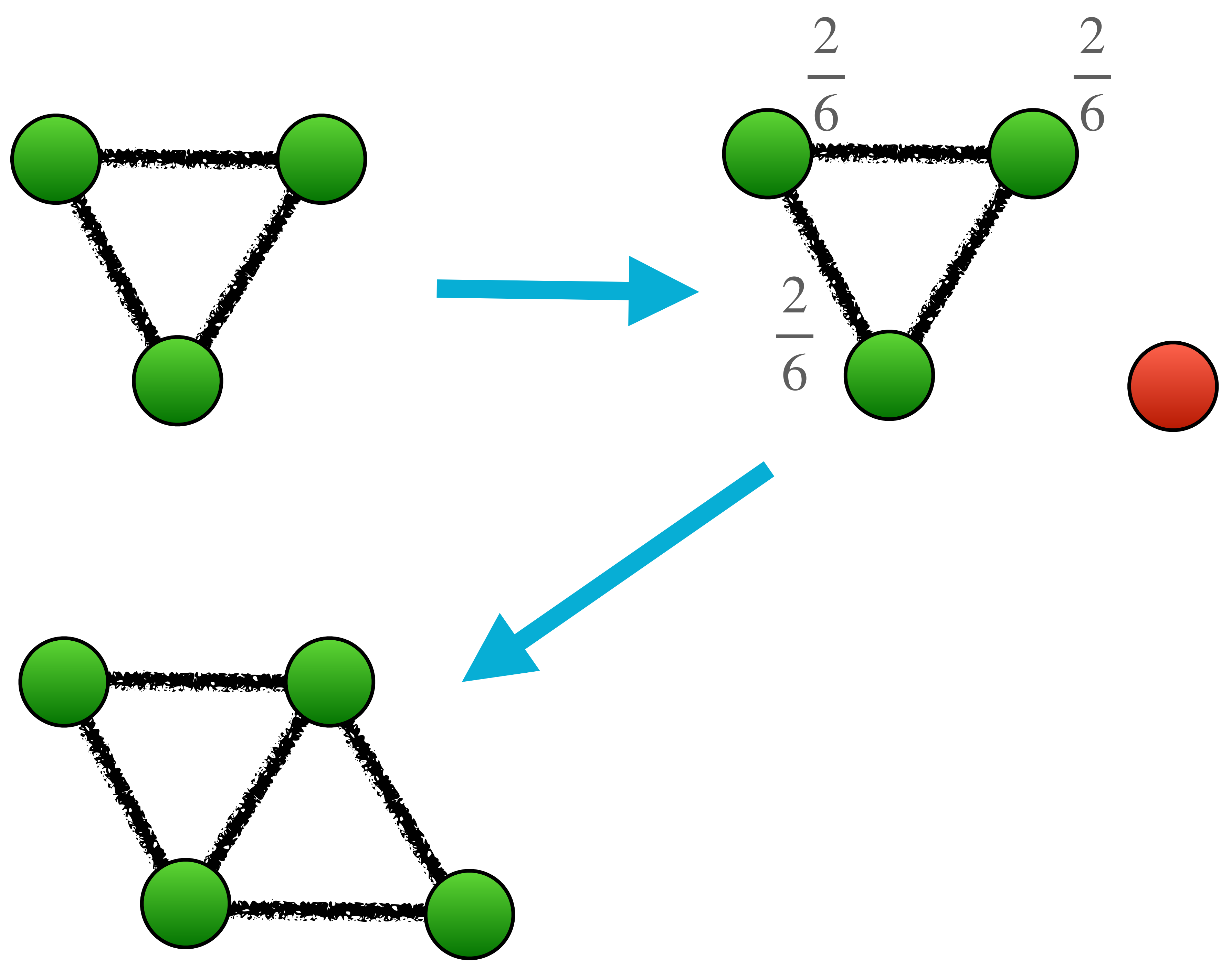
1. Start off with a small seed network
2. Choose *m* existing nodes with a probability directly proportional to their degree: $\frac{k_i}{\sum_j k_j}$
3. Connect a new node to the m chosen nodes
4. If not yeah *N* nodes, go to 2.



Preferential attachment

- In preferential attachment models new links connect preferentially to high degree nodes
- Example: Barabási-Albert model:

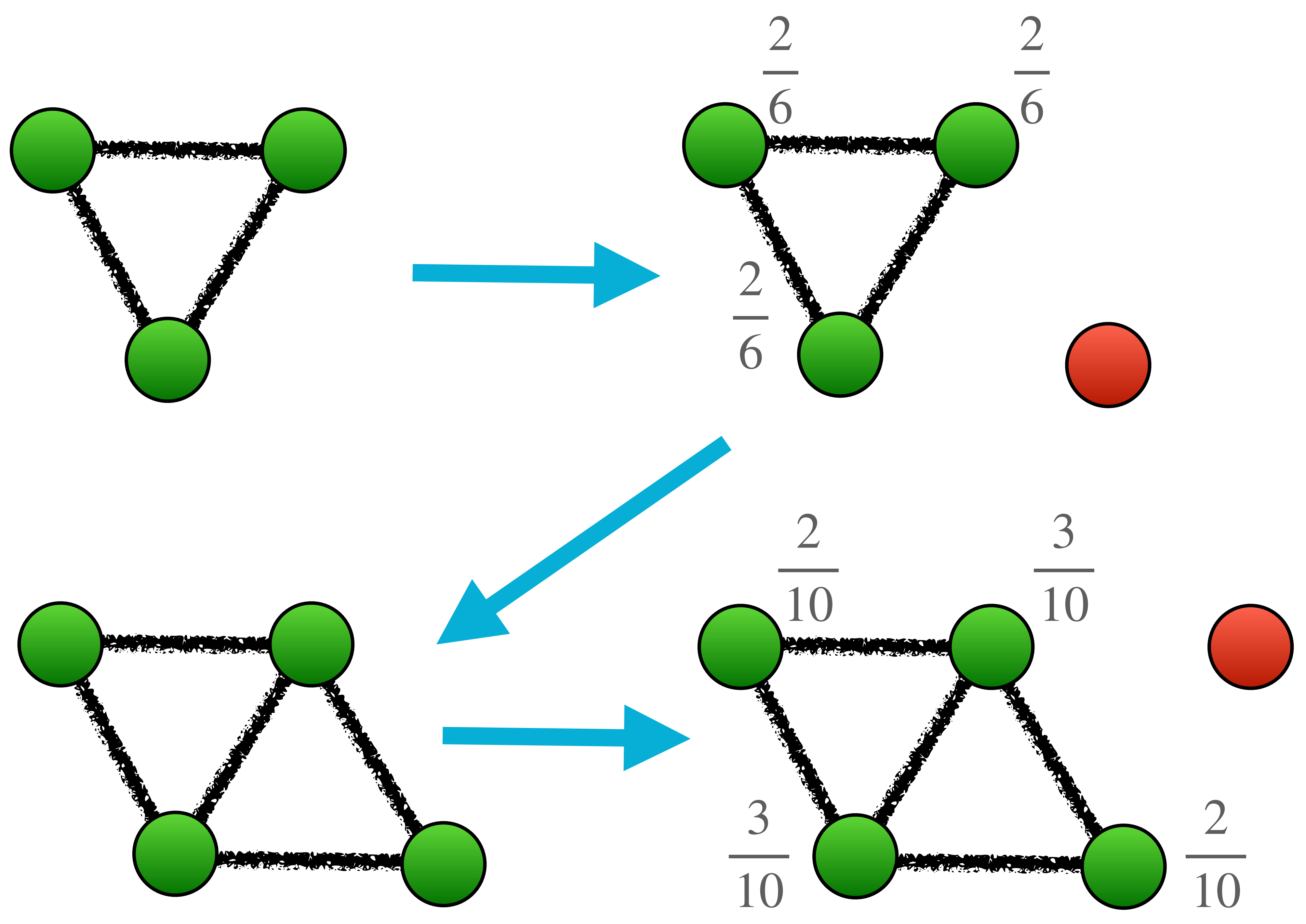
1. Start off with a small seed network
2. Choose *m* existing nodes with a probability directly proportional to their degree: $\frac{k_i}{\sum_j k_j}$
3. Connect a new node to the m chosen nodes
4. If not yeah *N* nodes, go to 2.



Preferential attachment

- In preferential attachment models new links connect preferentially to high degree nodes
- Example: Barabási-Albert model:

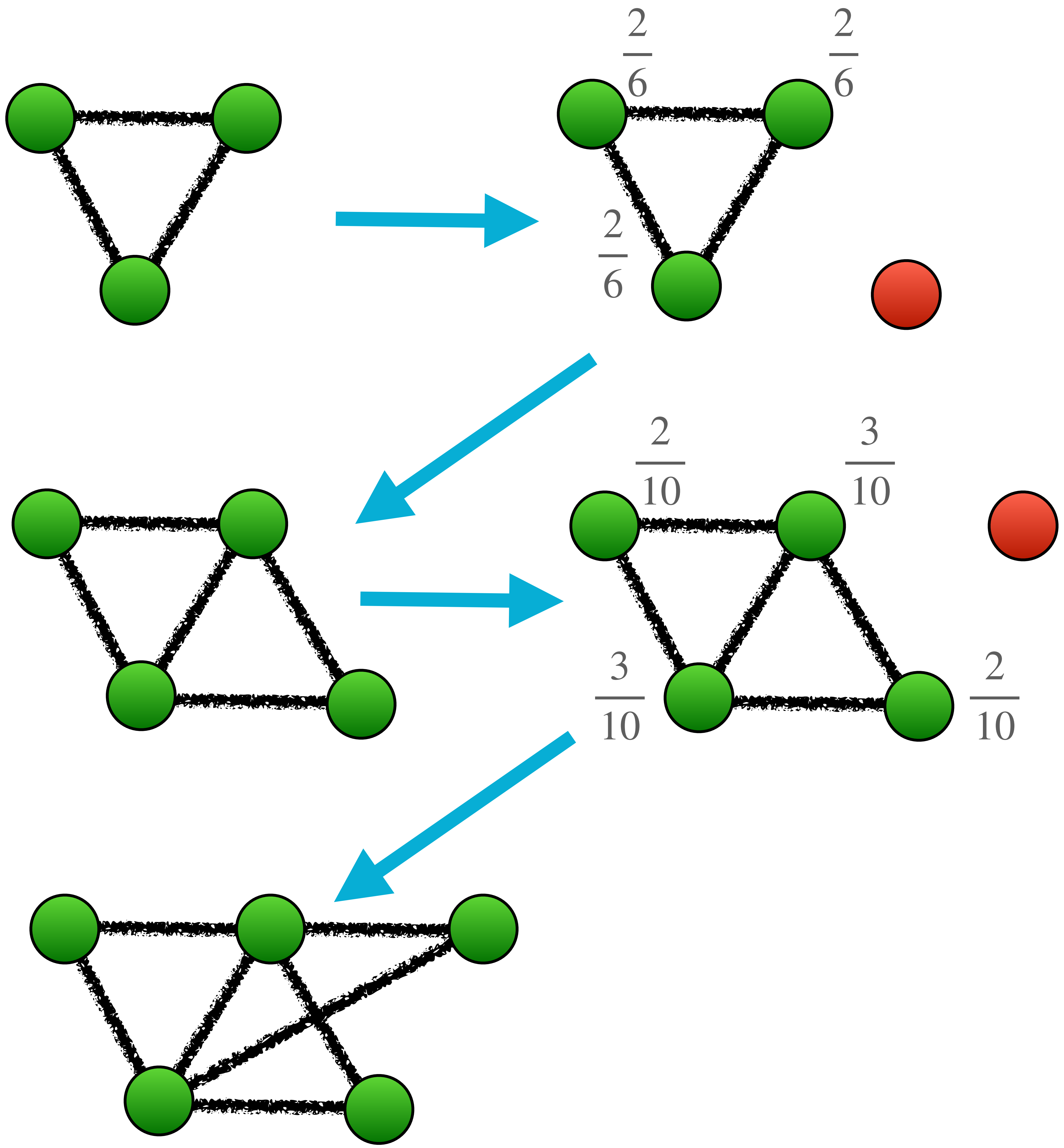
1. Start off with a small seed network
2. Choose m existing nodes with a probability directly proportional to their degree: $\frac{k_i}{\sum_j k_j}$
3. Connect a new node to the m chosen nodes
4. If not yeah N nodes, go to 2.



Preferential attachment

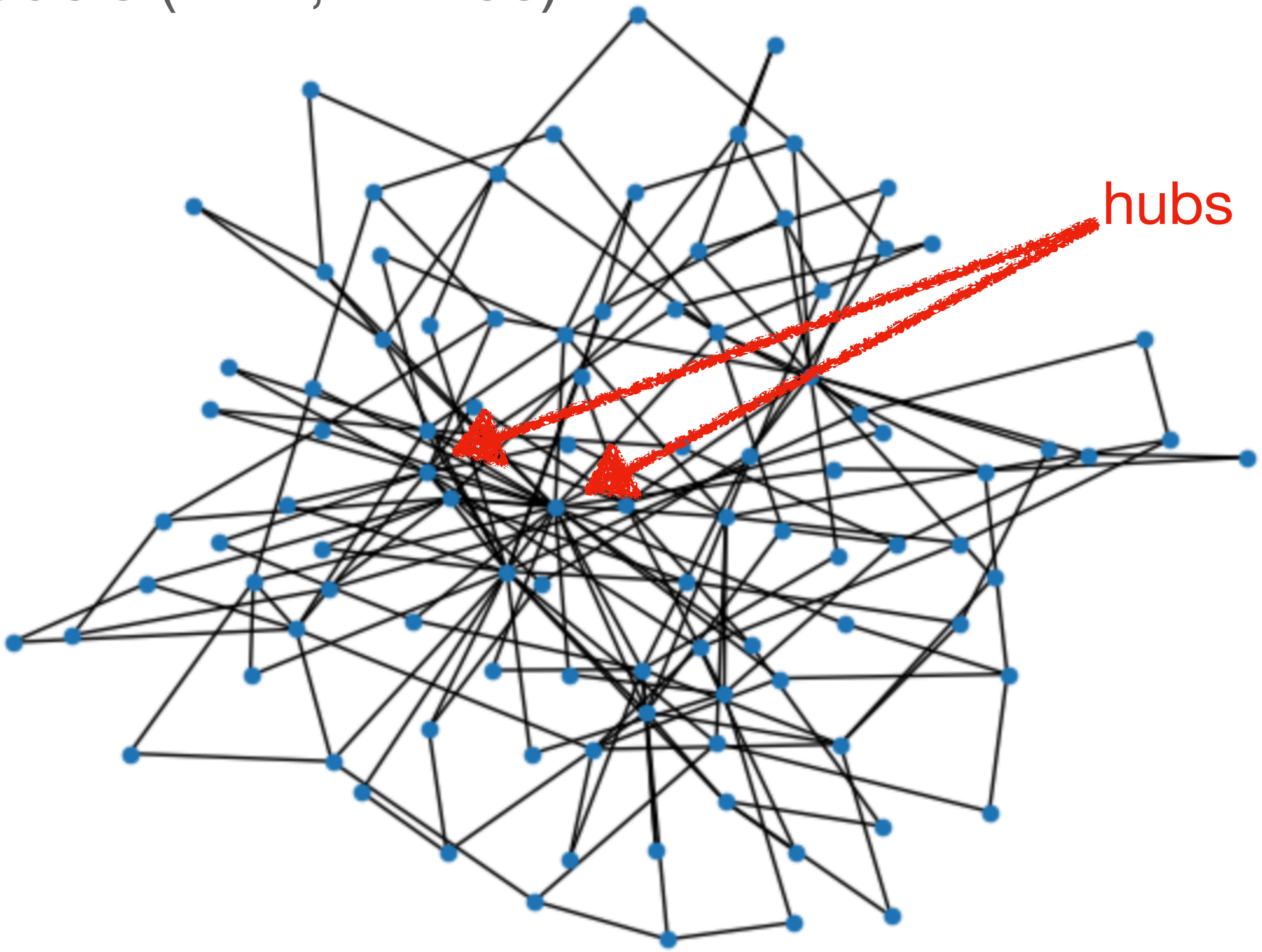
- In preferential attachment models new links connect preferentially to high degree nodes
- Example: Barabási-Albert model:

1. Start off with a small seed network
2. Choose *m* existing nodes with a probability directly proportional to their degree: $\frac{k_i}{\sum_j k_j}$
3. Connect a new node to the m chosen nodes
4. If not yeah *N* nodes, go to 2.



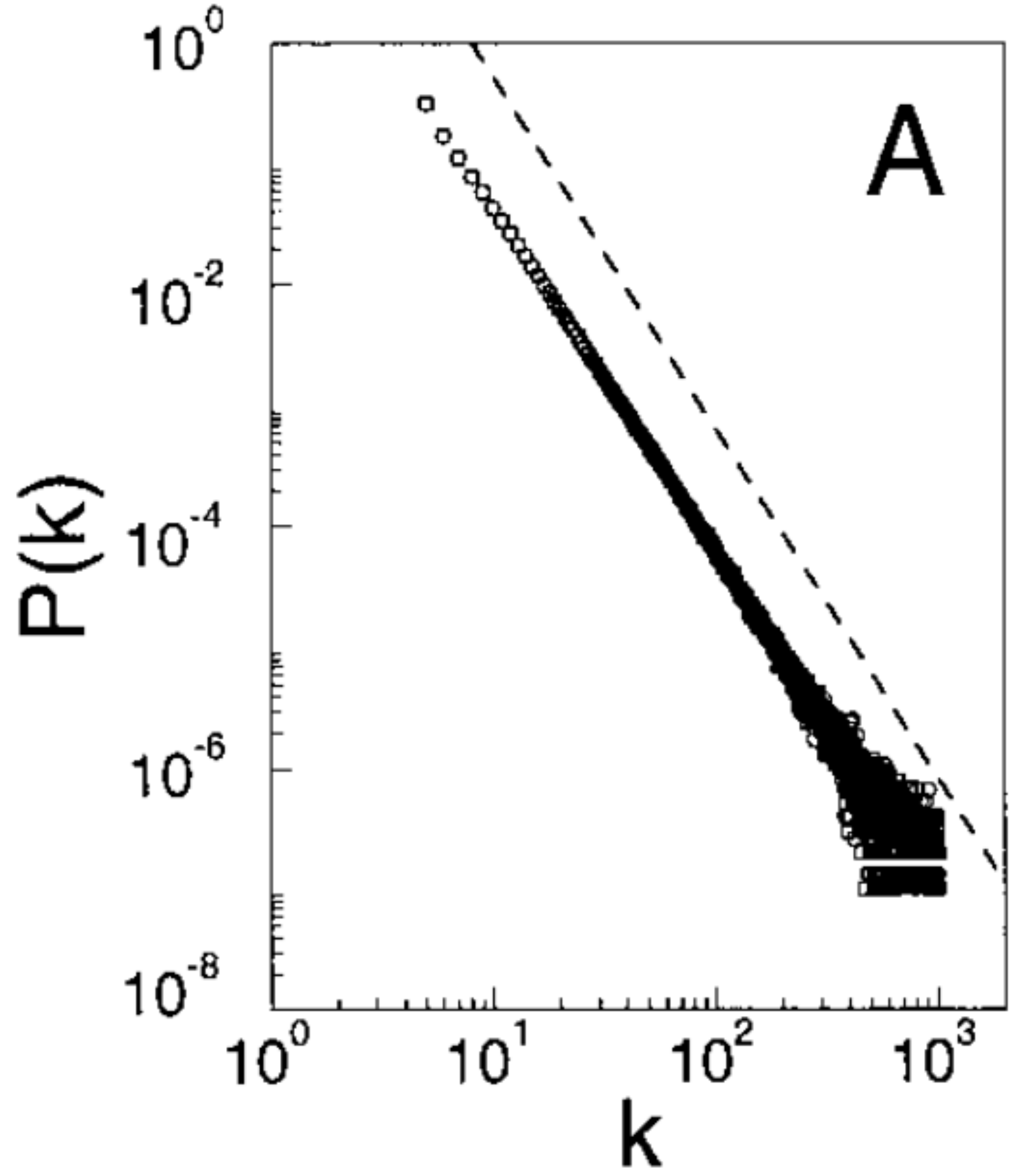
Result: power-law degree distributions

Barabási-Albert models (m=2, N=100):



Result: power-law degree distributions

- It can be shown formally that BA model will produce power-law degree distribution
- Similar results for more complex mechanism that have (implicit) preferential attachment (e.g., triadic closure)



Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. science, 286(5439), 509-512.

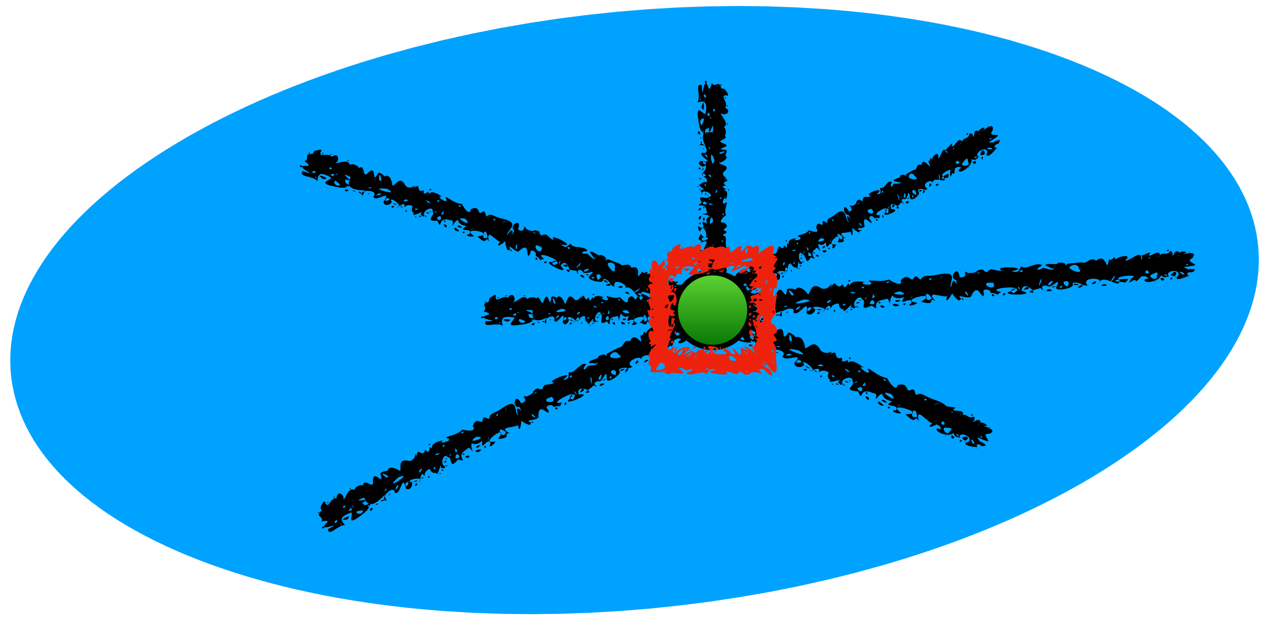
Dynamics → network structure

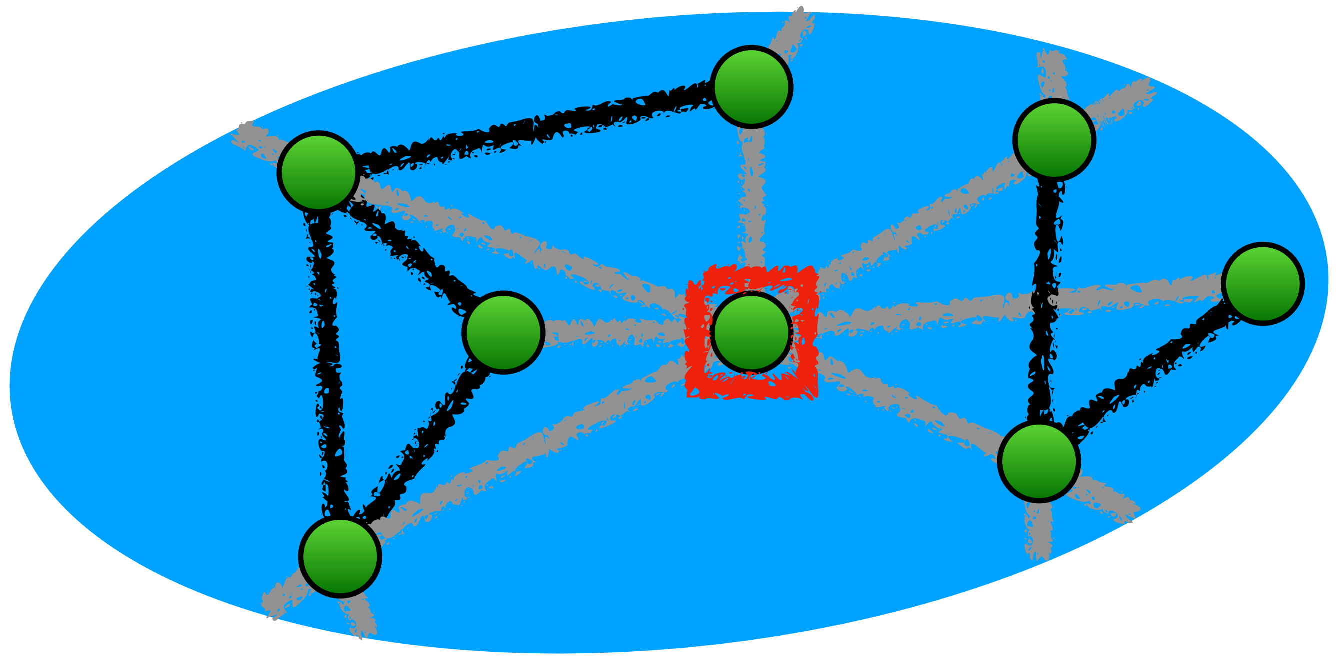
Network dynamics

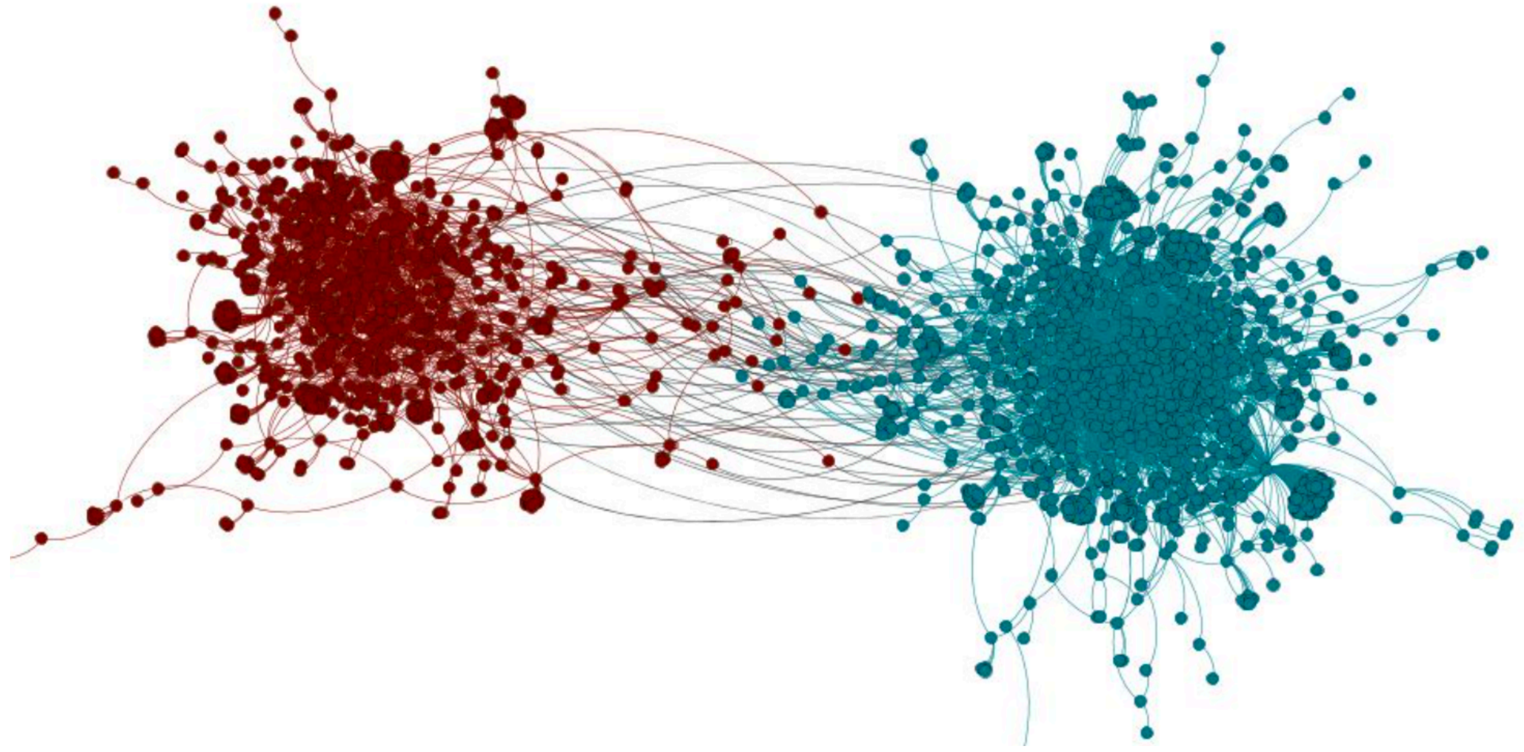
Degree heterogeneity

High clustering coefficients

Communities







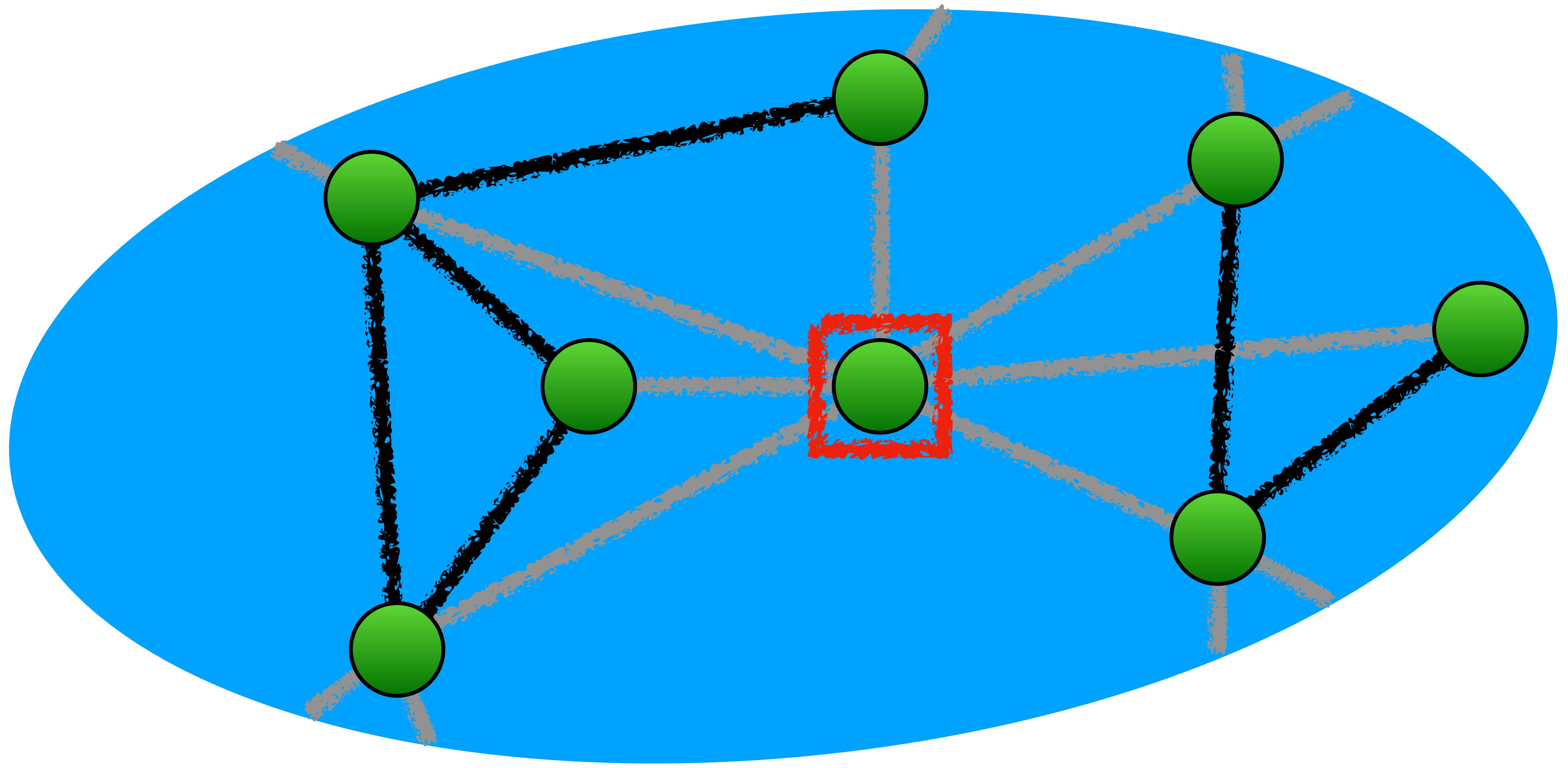
Reminder: clustering coefficient

- Clustering coefficient = density of the neighbourhood = # links between neighbouring nodes / # pairs of neighbours

- = # triangles / # of two-stars

- Social networks have high clustering coefficients:

    - Common friends, social groups, meeting people through existing connections…



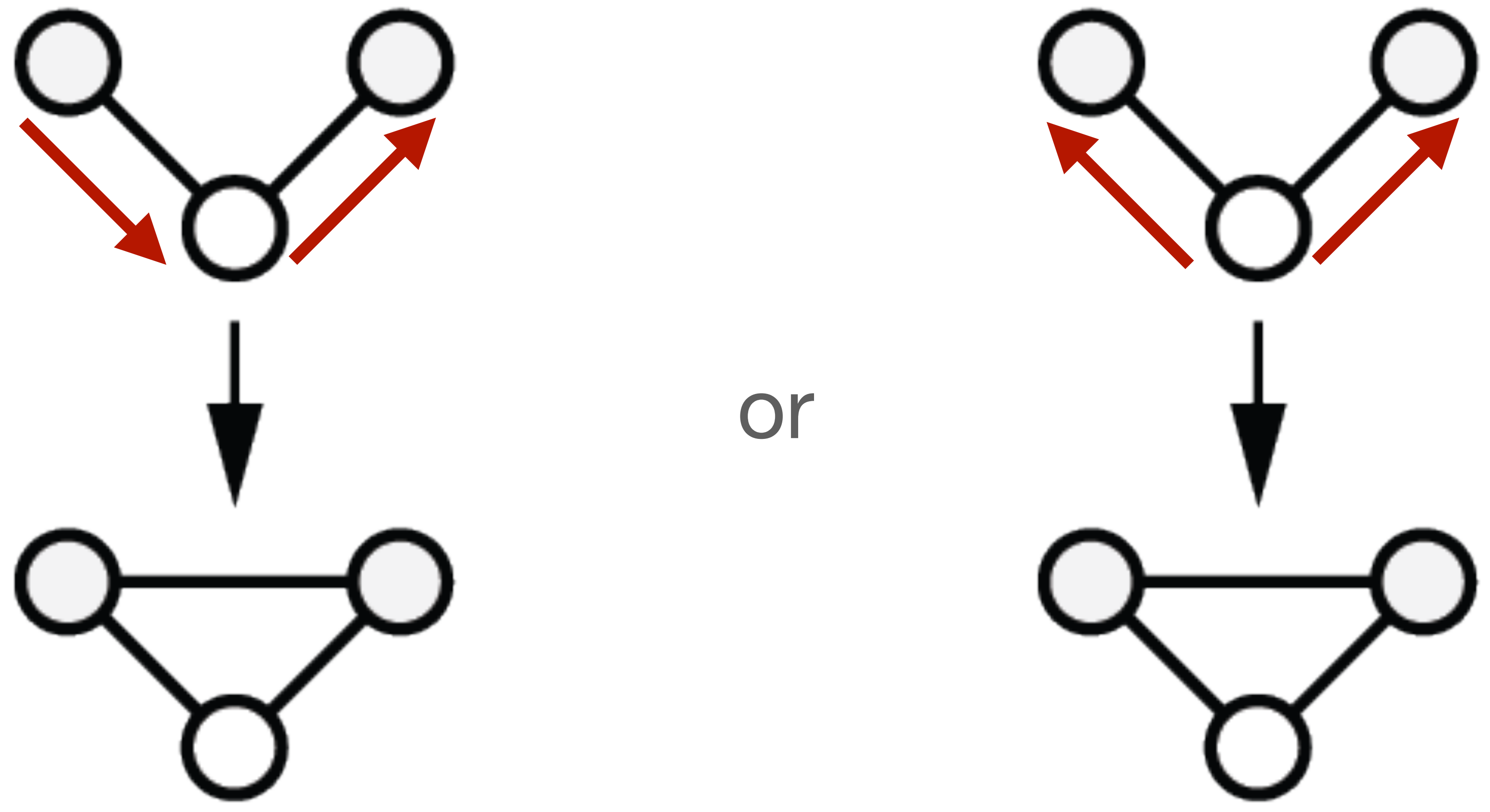
$$c_i = \frac{E_i}{\binom{k_i}{2}}$$

$$E_i = 6 \qquad \binom{k_i}{2} = 7 \times 6/2 = 21$$

$$c_i = \frac{6}{21} \approx 0.29$$

Creating new links

- *Social networks*:
    - You meet people based on your previous connections → **triadic closure**
        - Empirical evidence (e.g. Kossinets 2006)
        - Network models (e.g., Toivonen 2008)
- *Linked documents* (scientific articles, wikipedia pages, etc):
    - Find document based on the network and link to them
    - Create (partial) copies of old documents (or links in them)



Triadic closure → triangles → high clustering

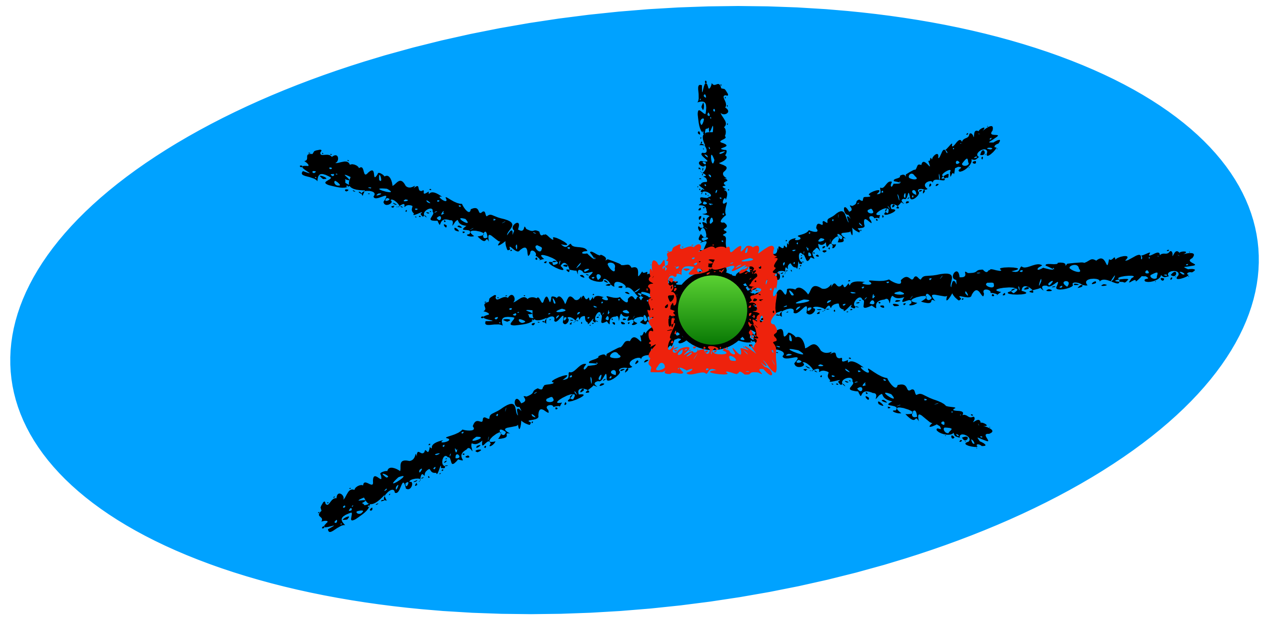
Dynamics → network structure

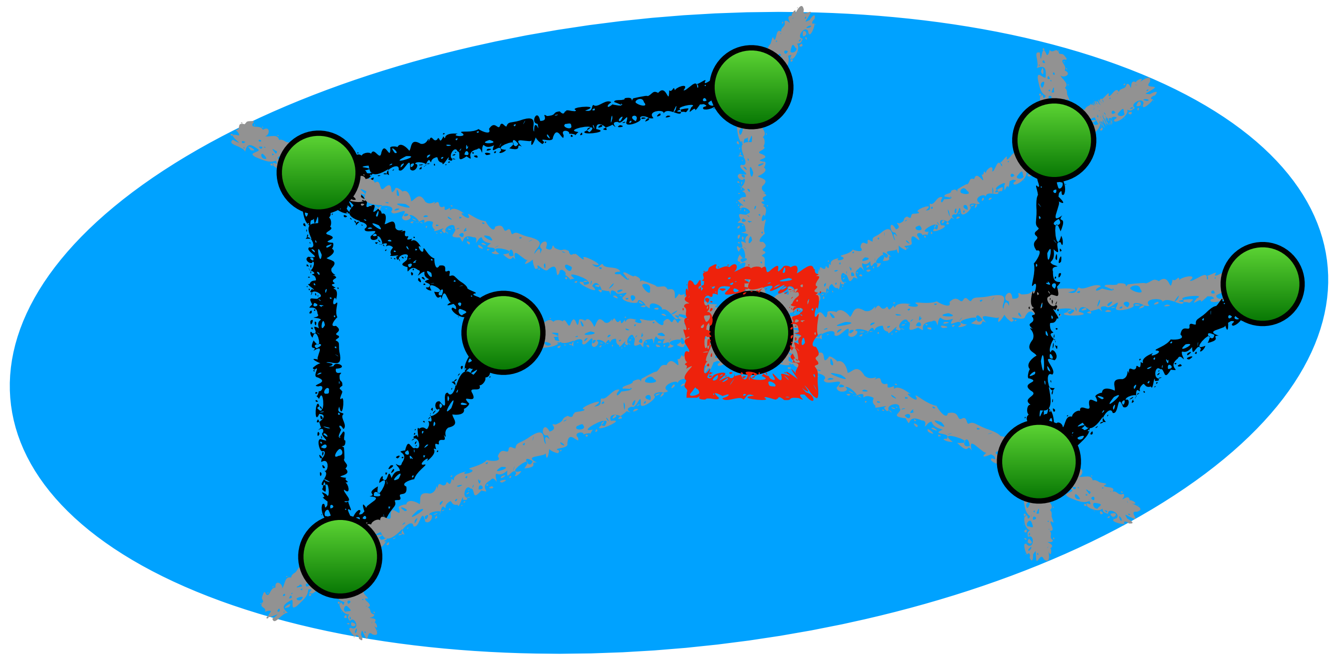
Network dynamics

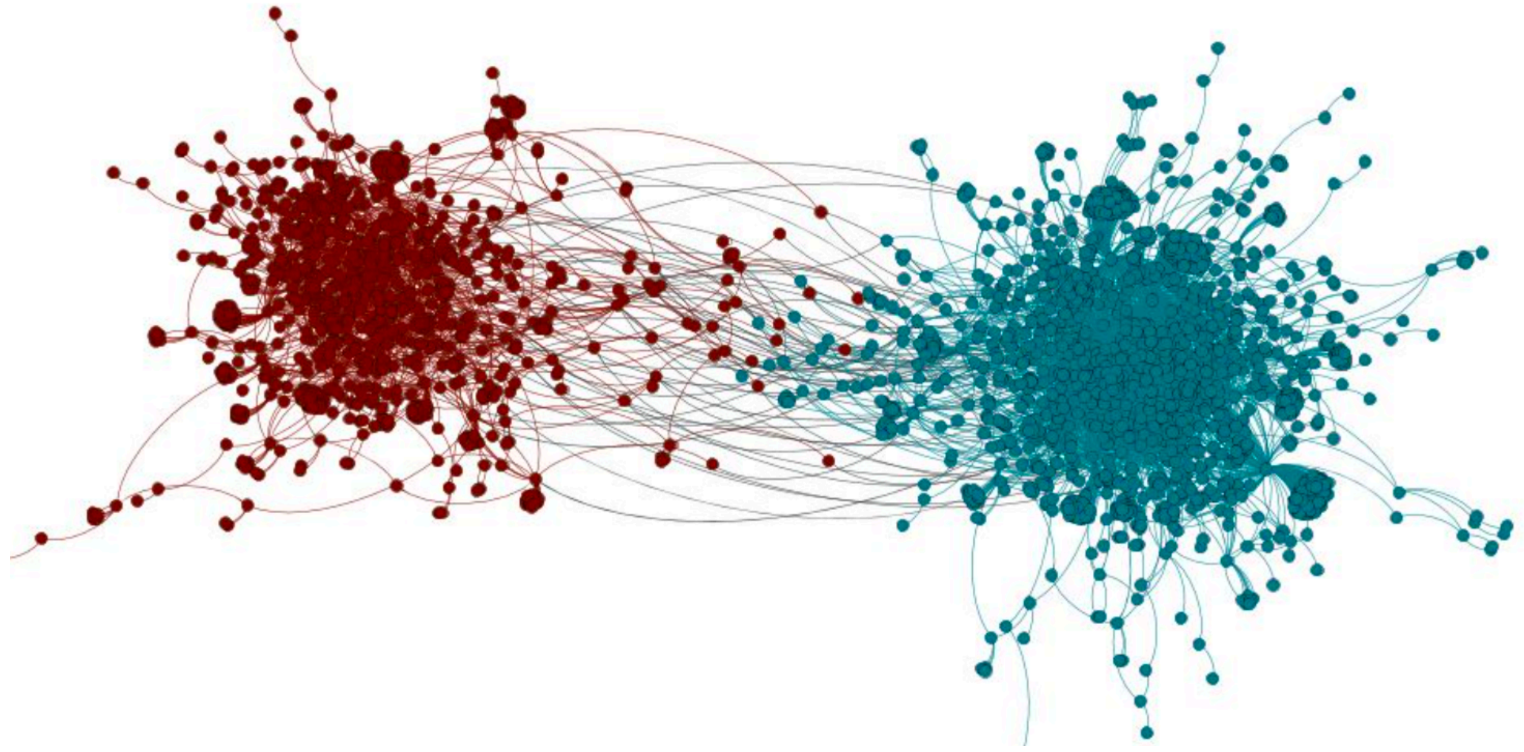
Degree heterogeneity

High clustering coefficients

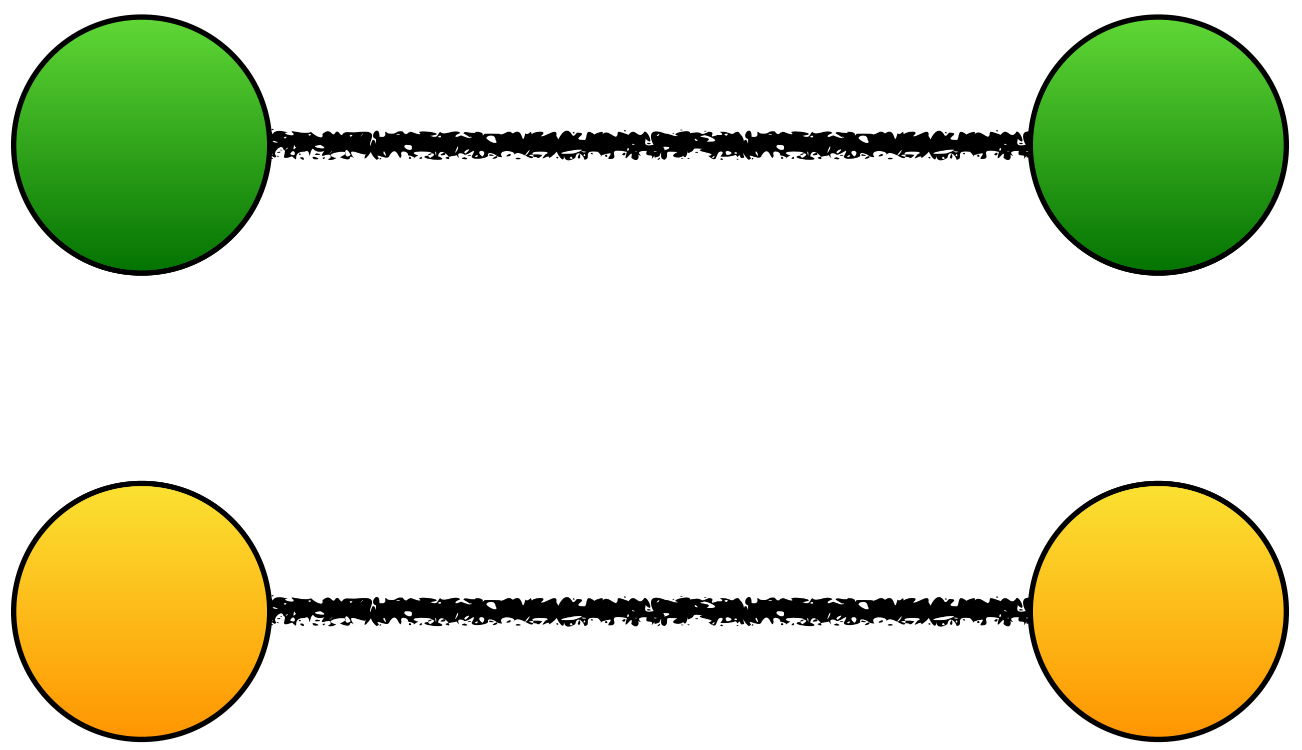
Communities







Reminder: homophily



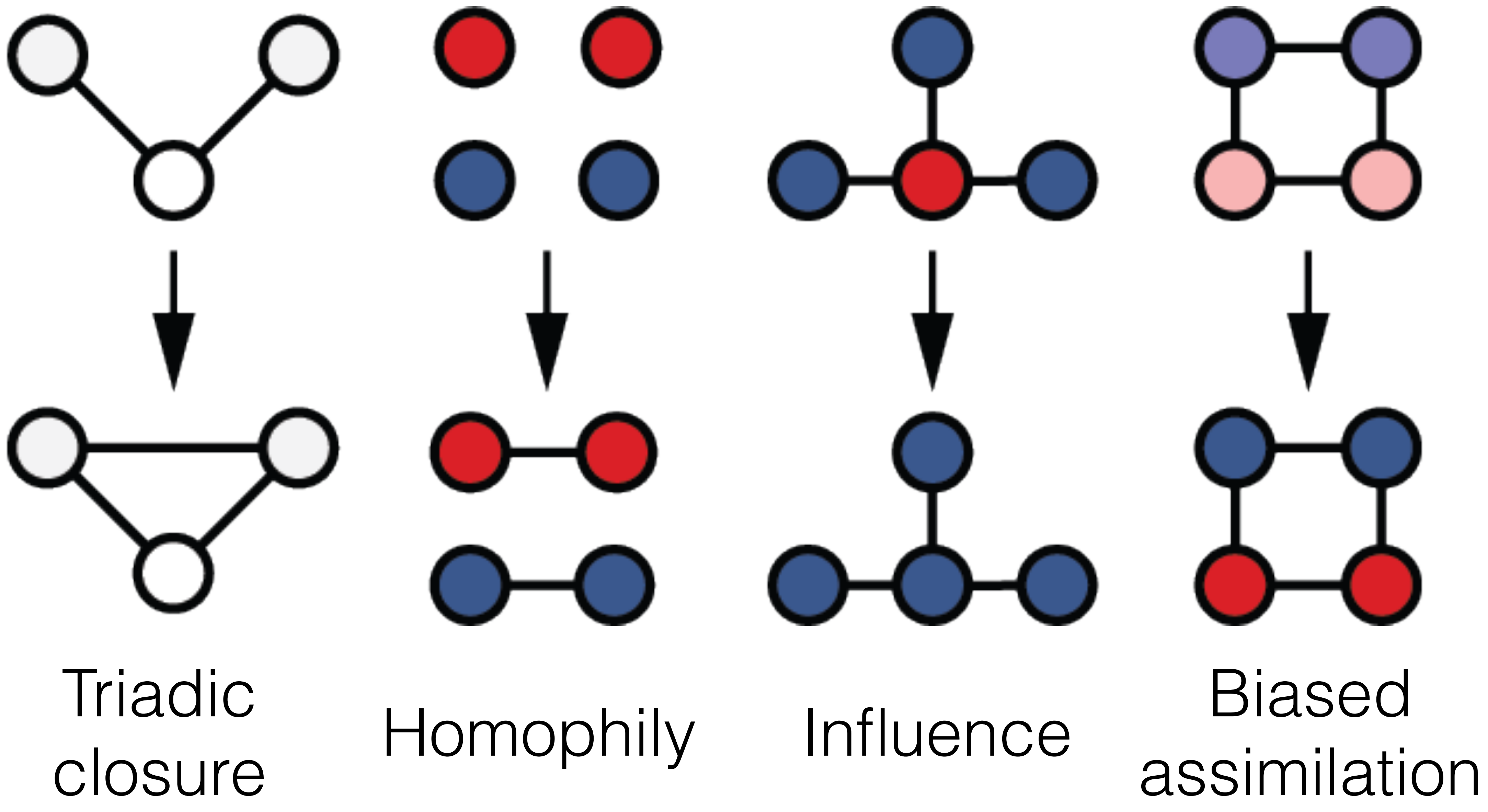
- Similar people tend to be connected to each other
- "Birds of the feather flock together"
- Observed in most social networks

"Tackling grand challenges" @AaltoUniversity, Twitter

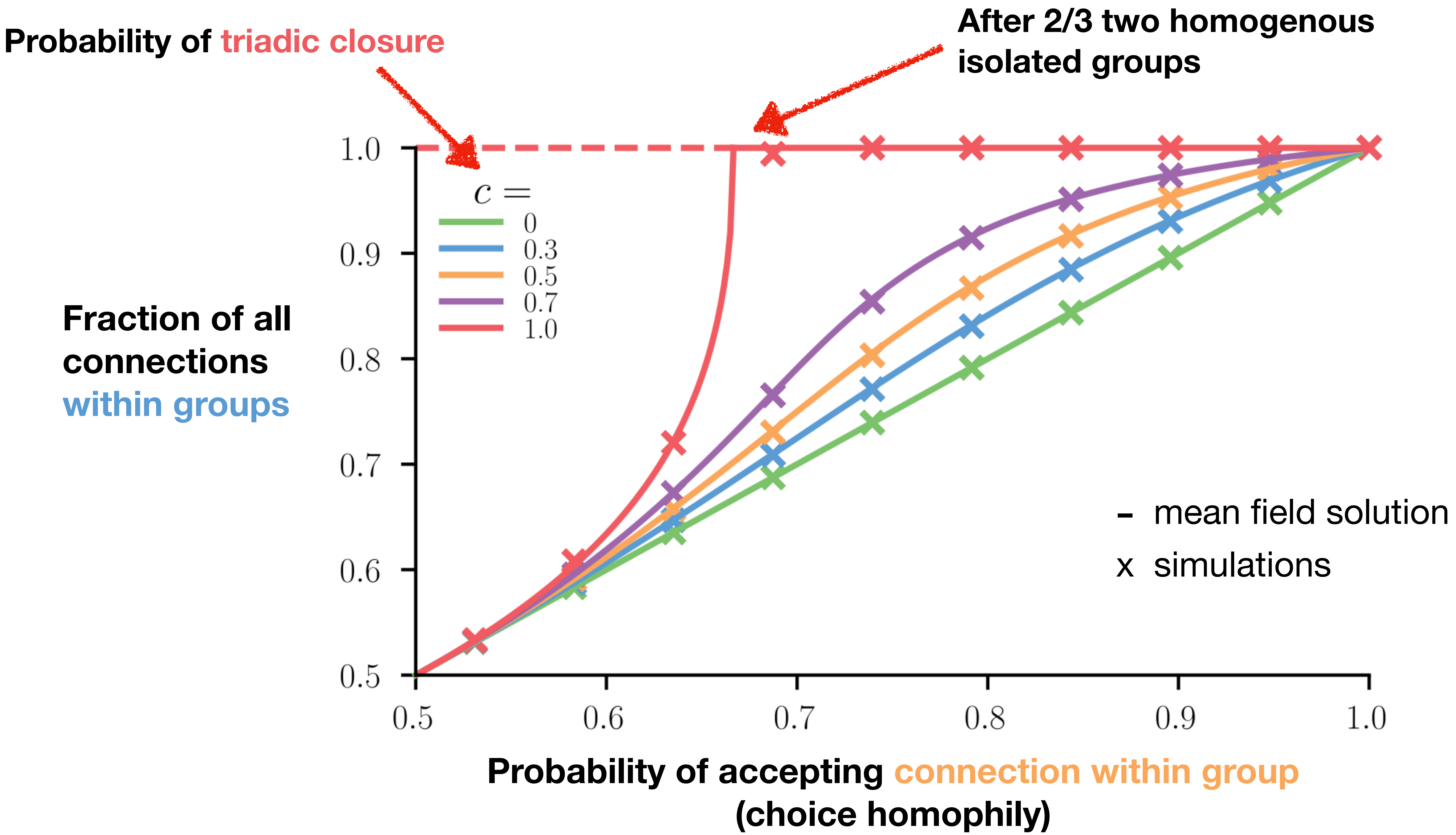
M McPherson, L Smith-Lovin, JM Cook "Birds of a Feather: Homophily in Social Networks", Annual Review of Sociology Vol. 27:415-444 (2001)

Emergence of communities

- Clusters/communities with homogeneous attributes: combination of mechanisms, e.g.:
  - Homophily + influence
  - Triadic closure + homophily



Emergence of communities



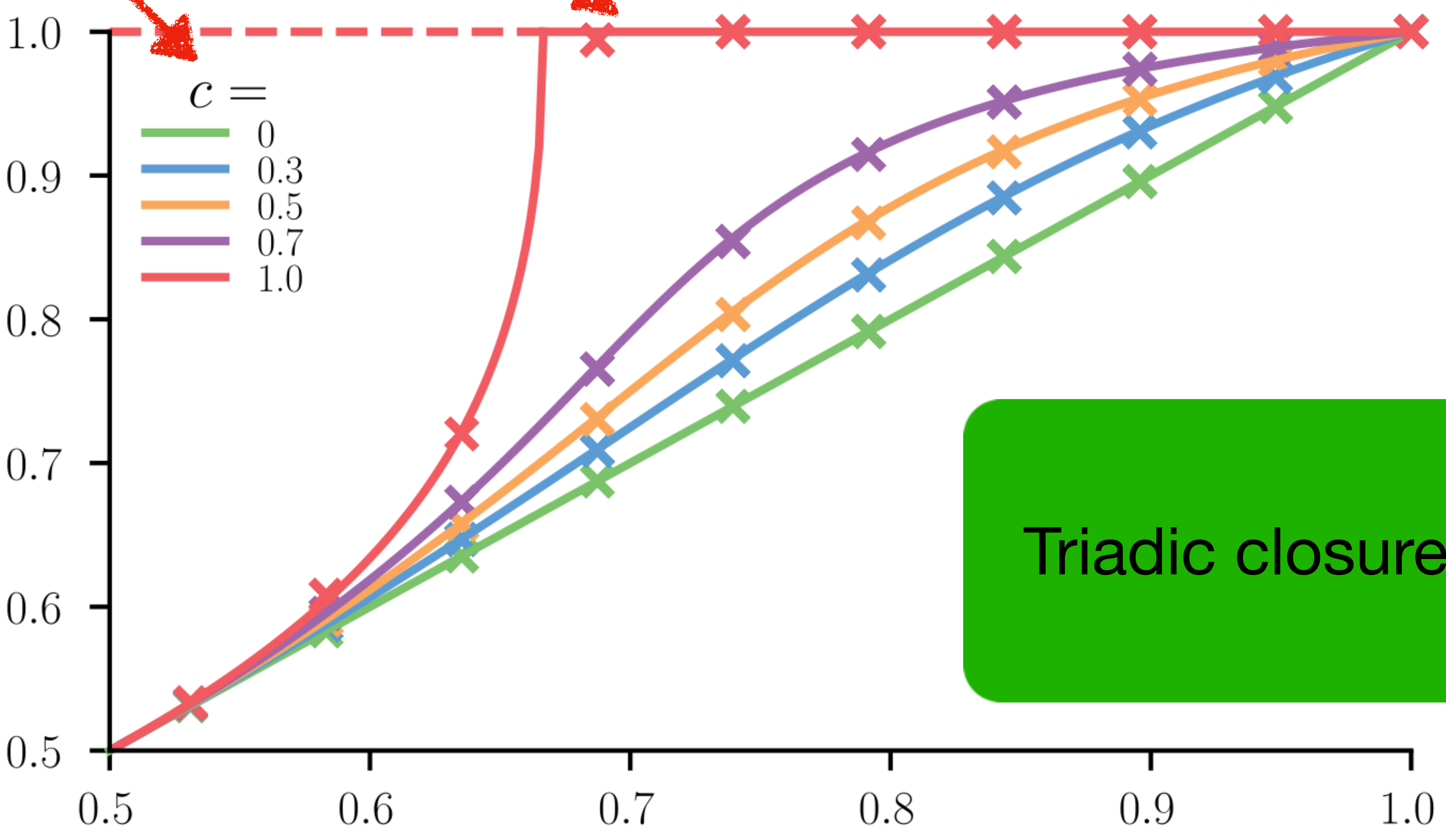
A Asikainen, G Iniguez, J Urena-Carrion, K Kaski, M Kivelä. Science Advances 2020: Vol. 6, no. 19, eaax7310

Emergence of communities

Probability of triadic closure

After 2/3 two homogenous isolated groups

Fraction of all connections within groups



Triadic closure amplifies homophily

Probability of accepting connection within group (choice homophily)

A Asikainen, G Iniguez, J Urena-Carrion, K Kaski, M Kivelä. Science Advances 2020: Vol. 6, no. 19, eaax7310

Summary: network dynamics

- Many common social network structural features can be explained by simple network evolution mechanisms
  - Direct consequences: triadic closure leads to clustering
  - Emergent: preferential attachment leads to hubs, homophily is amplified by triadic closure, …

How does the structure affect dynamics?

Dynamics on networks

- We don’t often care about the structure of the network, but what happens on the network:
  - Information/disease spreading
  - Behaviour adaption (complex contagion)
  - Opinion dynamics
- How does the structure of the network affect the dynamics?

Influence

- Social mechanism: people change their opinions/behavior they are interacting with

- Models for adaption/spreading: SI, SIR, threshold models, …

- Models for competing behavior/opinions: voter models, Sznajd model, majority rule model, …

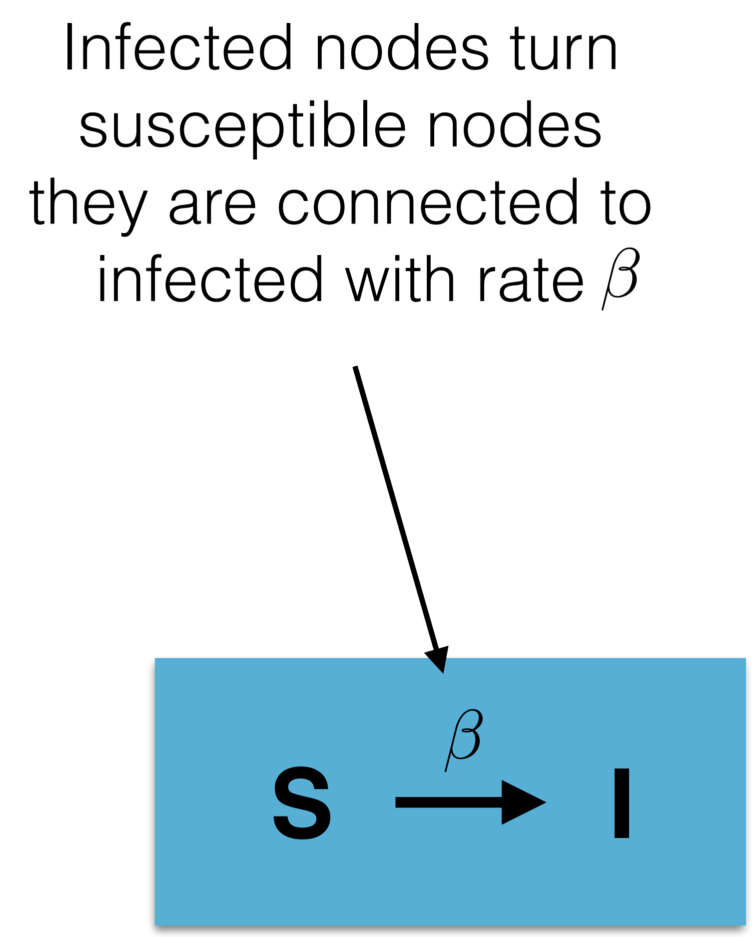
How can we model spreading dynamics?

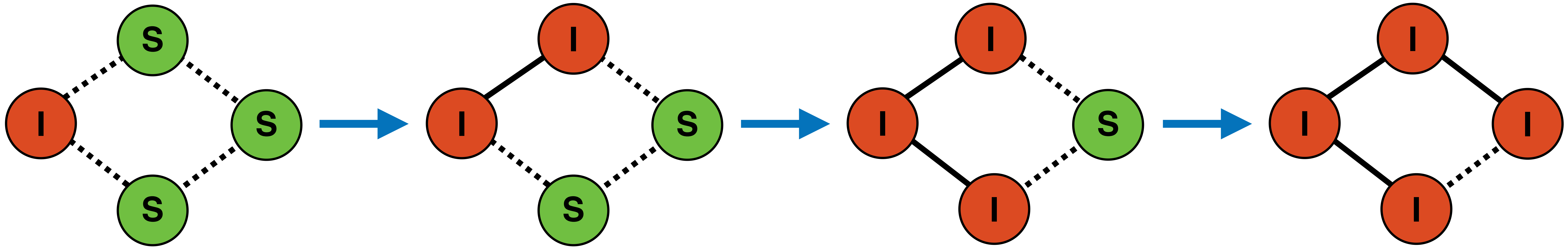
Epidemic models

- Epidemic models on networks follow the more conventional literature on mathematical epidemiology
- Can also be used to model information spreading
- Each node in one of the following compartments:
  - S: Susceptible
  - E: Exposed
  - I: Infected
  - R: Recovered (or Removed)
- Naming convention: states change from left to right
- Typical models: SI, SIS, SIR, SIRS, SEIR, …

Example: SI model

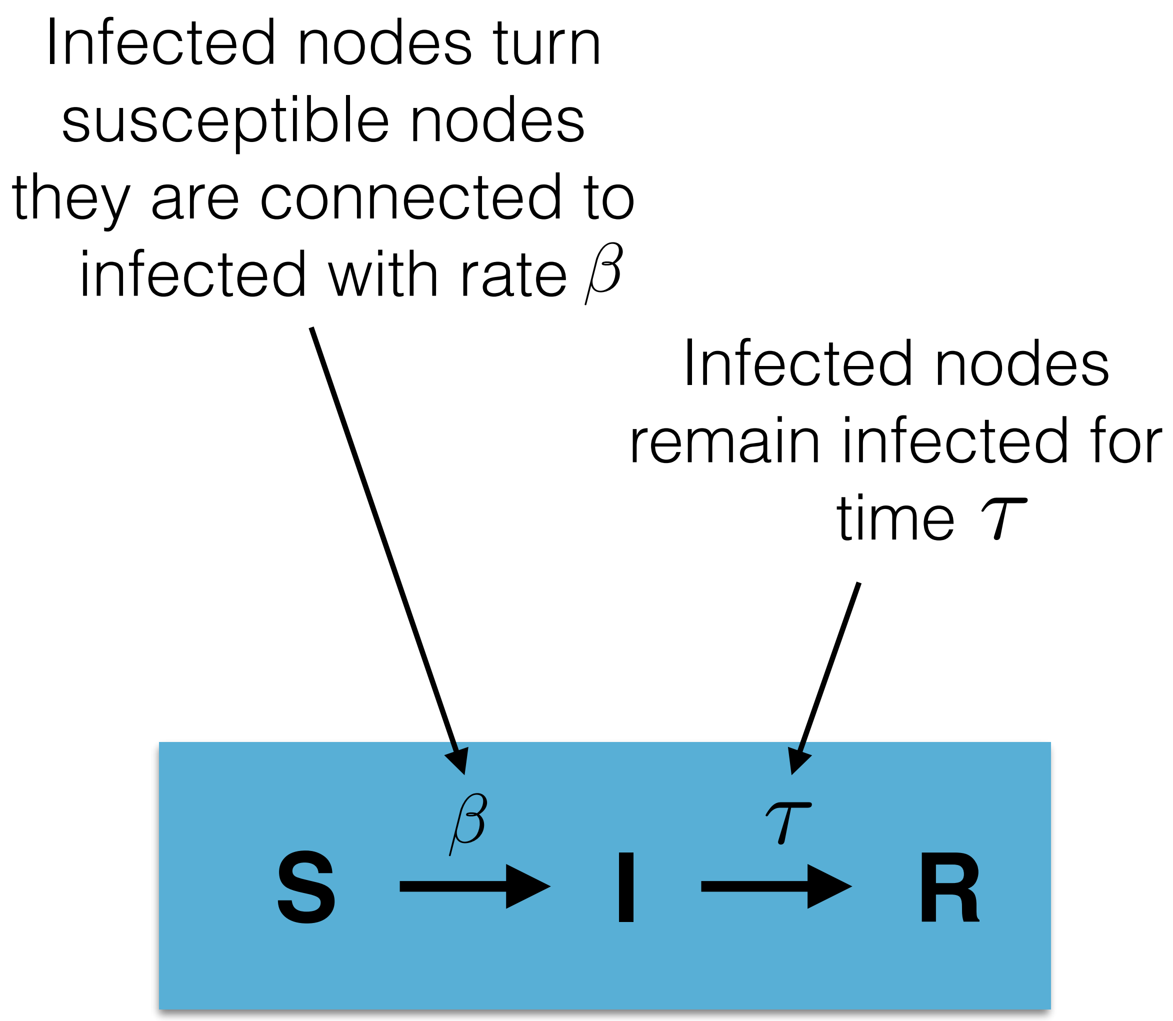
- Each node in one of 2 compartments:
  - Susceptible → Infected
- Set one node to compartment I, rest in compartment S
- In the end, if networks is connected, all the nodes are in compartment I

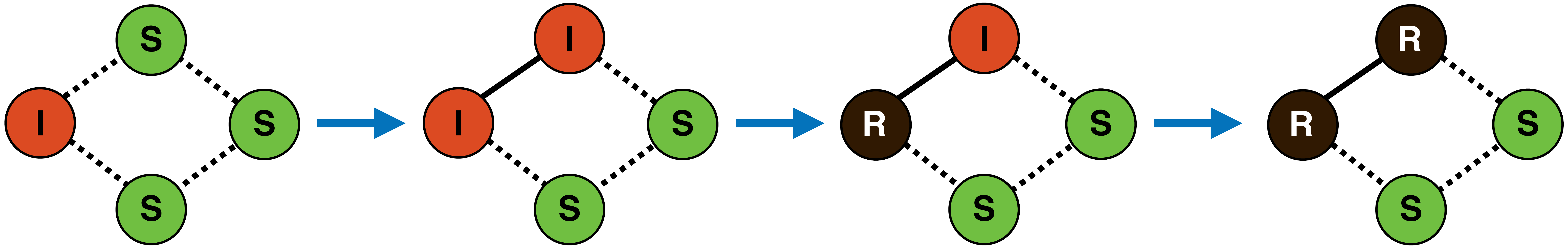




Example: SIR model

- Each node in one of 3 compartments
  - Susceptible → Infected → Recovered
- Set one node to compartment I, rest in compartment S
- In the end, all nodes in S or R compartment





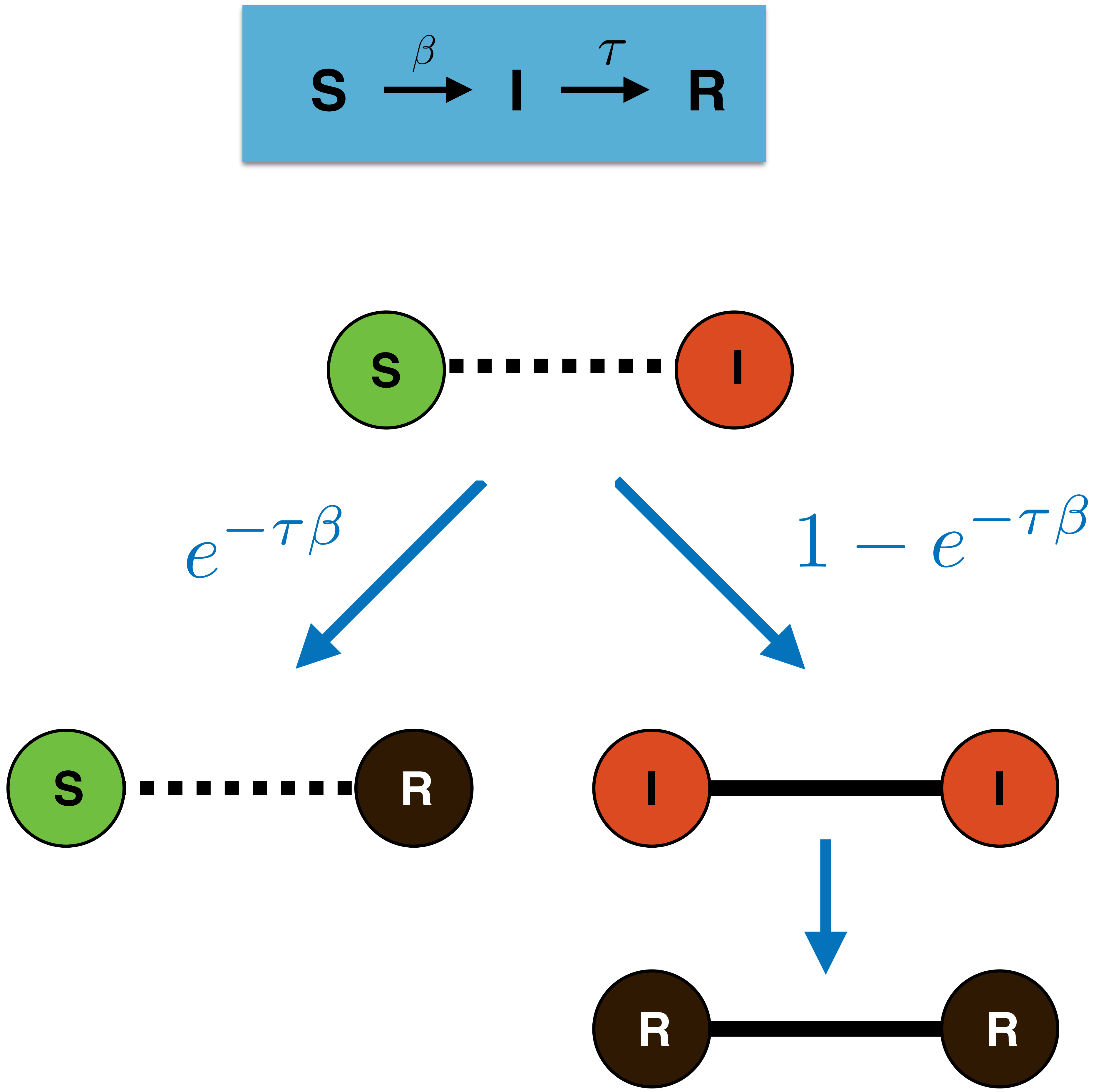
How the network structure affects the number of people getting an infection?

SIR spreading and R_0

- The *basic reproductive number* R_0 tells how many others does one infected person infect on average in the beginning of the epidemics
- For a single link, the probability of transmitting infection is $f = 1 - e^{-\tau\beta}$
- For a large completely random (ER) network where nodes have on average $\langle k \rangle$ connections:

$$R_0 = f\langle k \rangle$$

- $R_0 > 1 \rightarrow$ chance of an epidemic
- For other networks, this is not true!

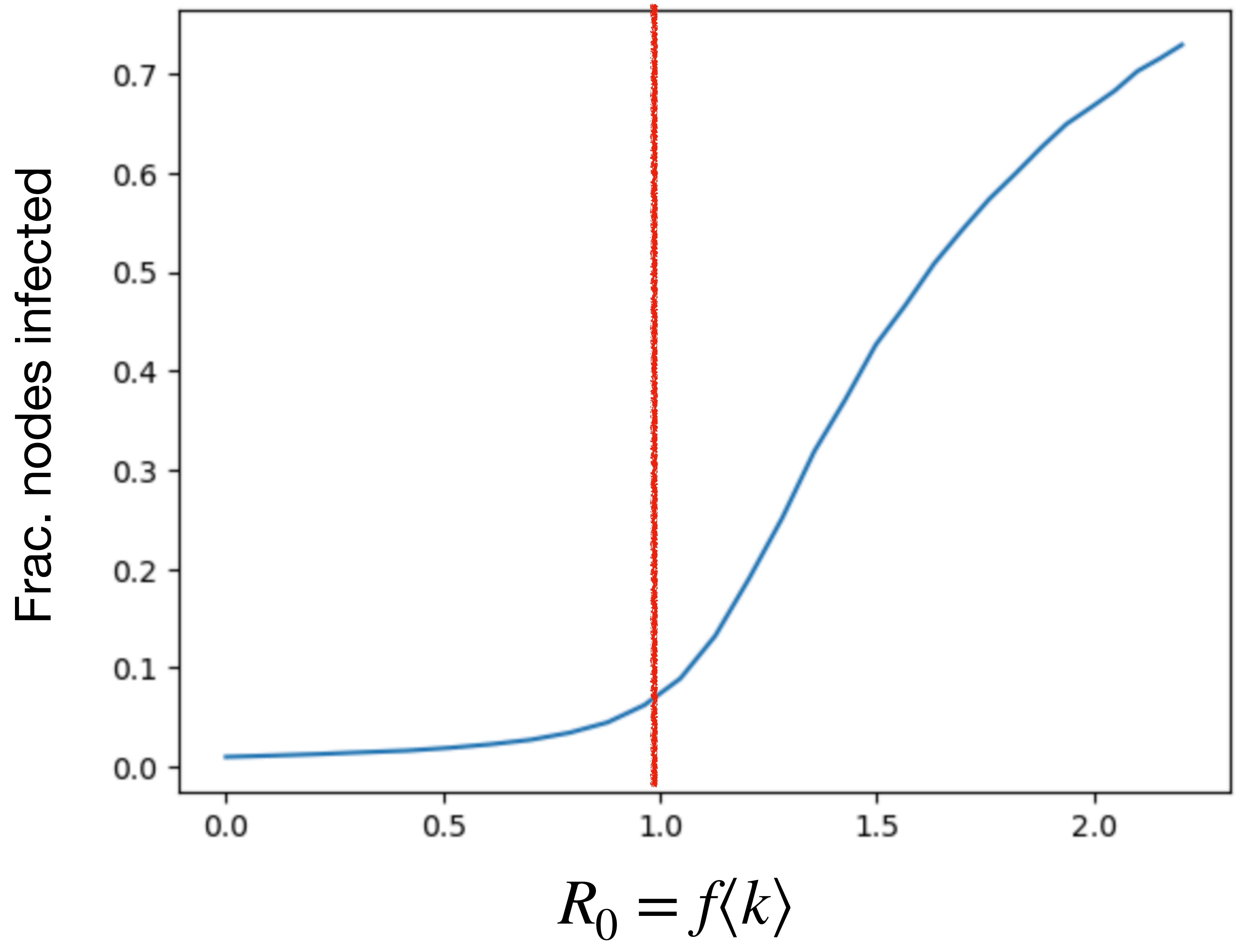


SIR spreading and phase transition

- Feedback loop:

  more infected people ⇆ more new infections

- Transition at $R_0 = 1$:
  - $R_0 < 1$: only small outbreaks
  - $R_0 = 1$: power-law size distribution of outbreaks
  - $R_0 > 1$: either large or small outbreak (no "medium" outbreaks sized)



Spreading and hubs

- Spreading = following links

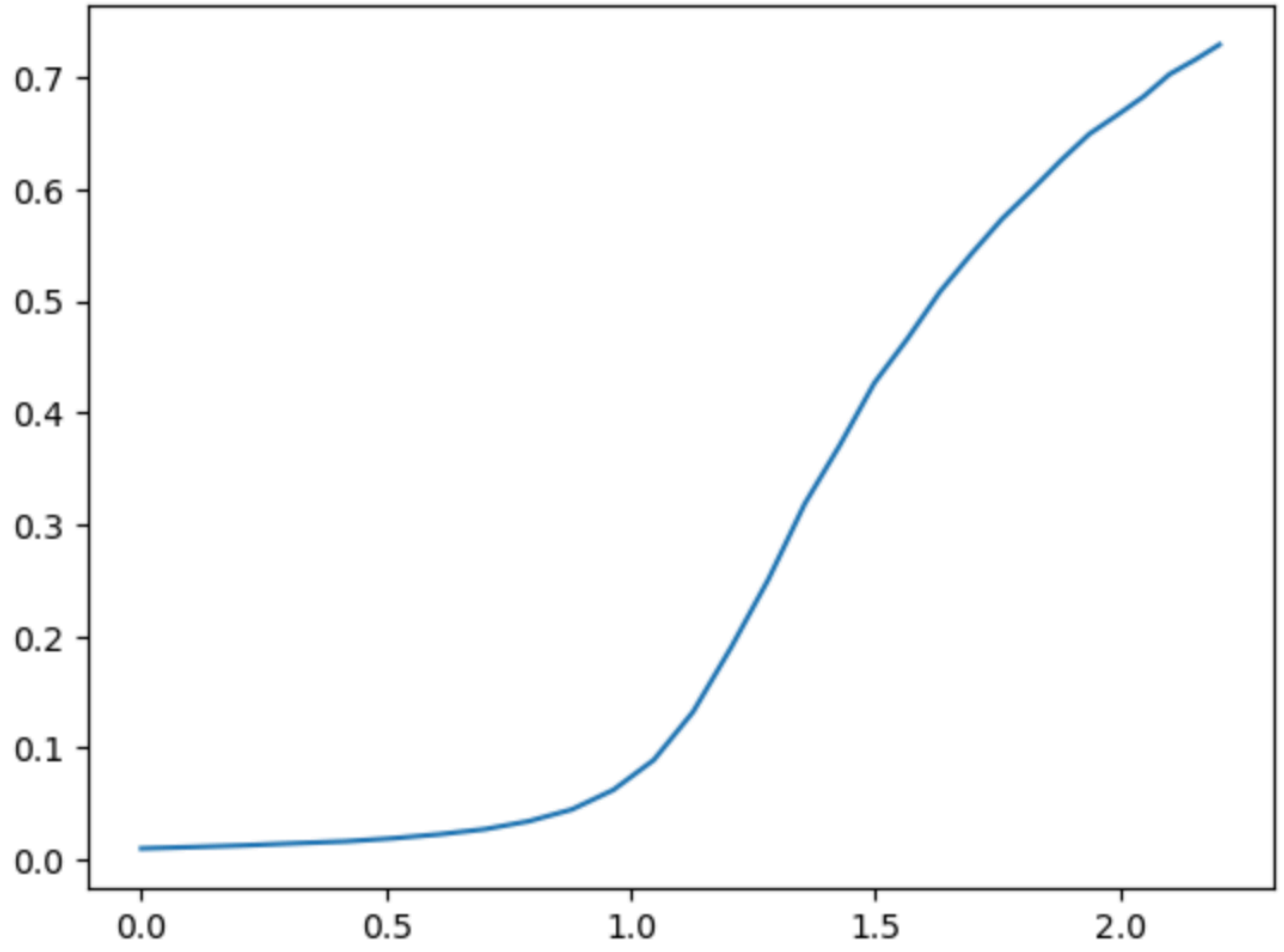
1. High degree nodes much more likely to get infected than low degree nodes

2. High degree nodes more efficient at spreading than low degree nodes

→ Degree distribution has a huge effect on any spreading process!

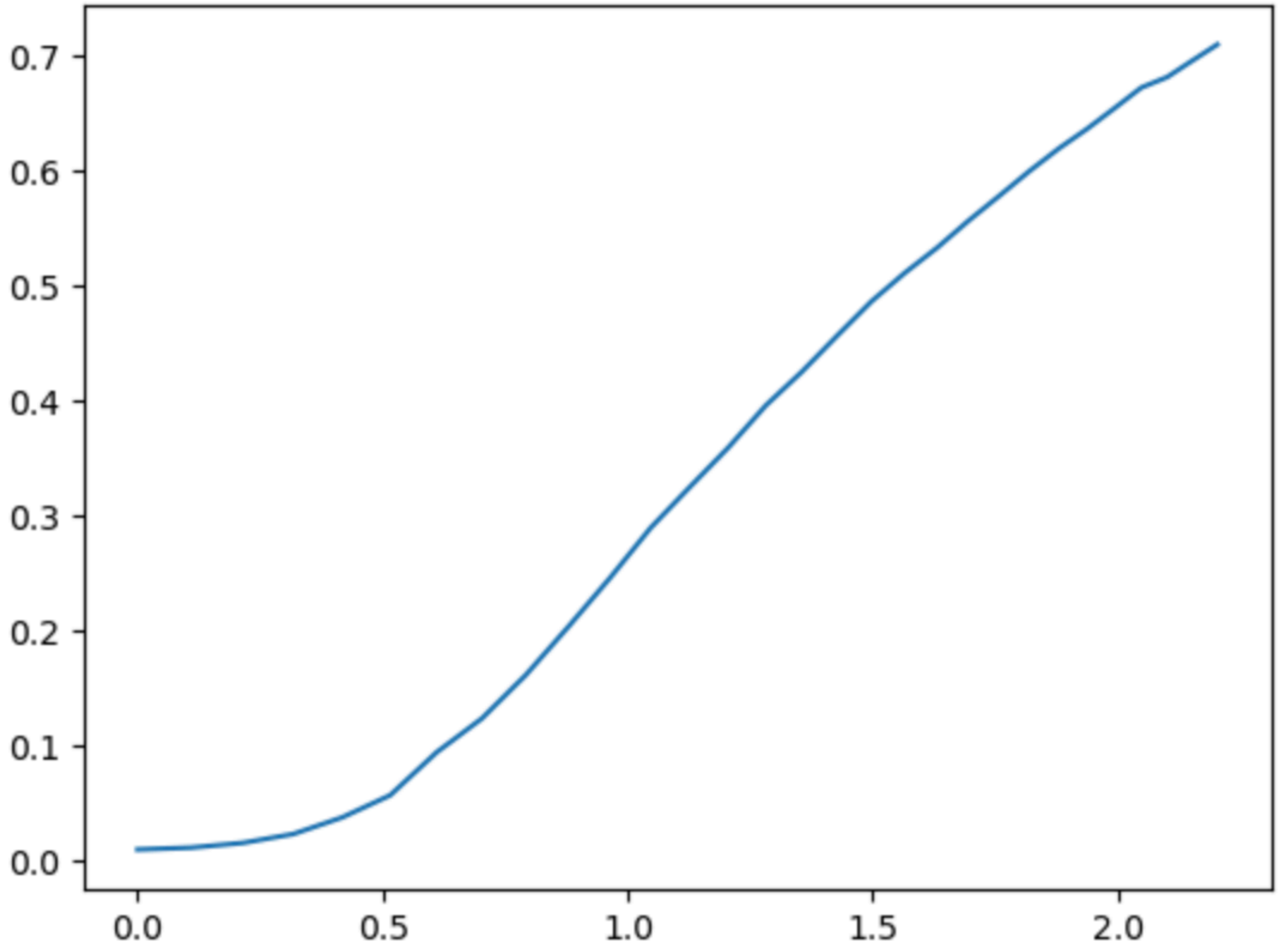
SIR spreading and phase transition

Fully random (ER) network



$R_0 = f\langle k \rangle$

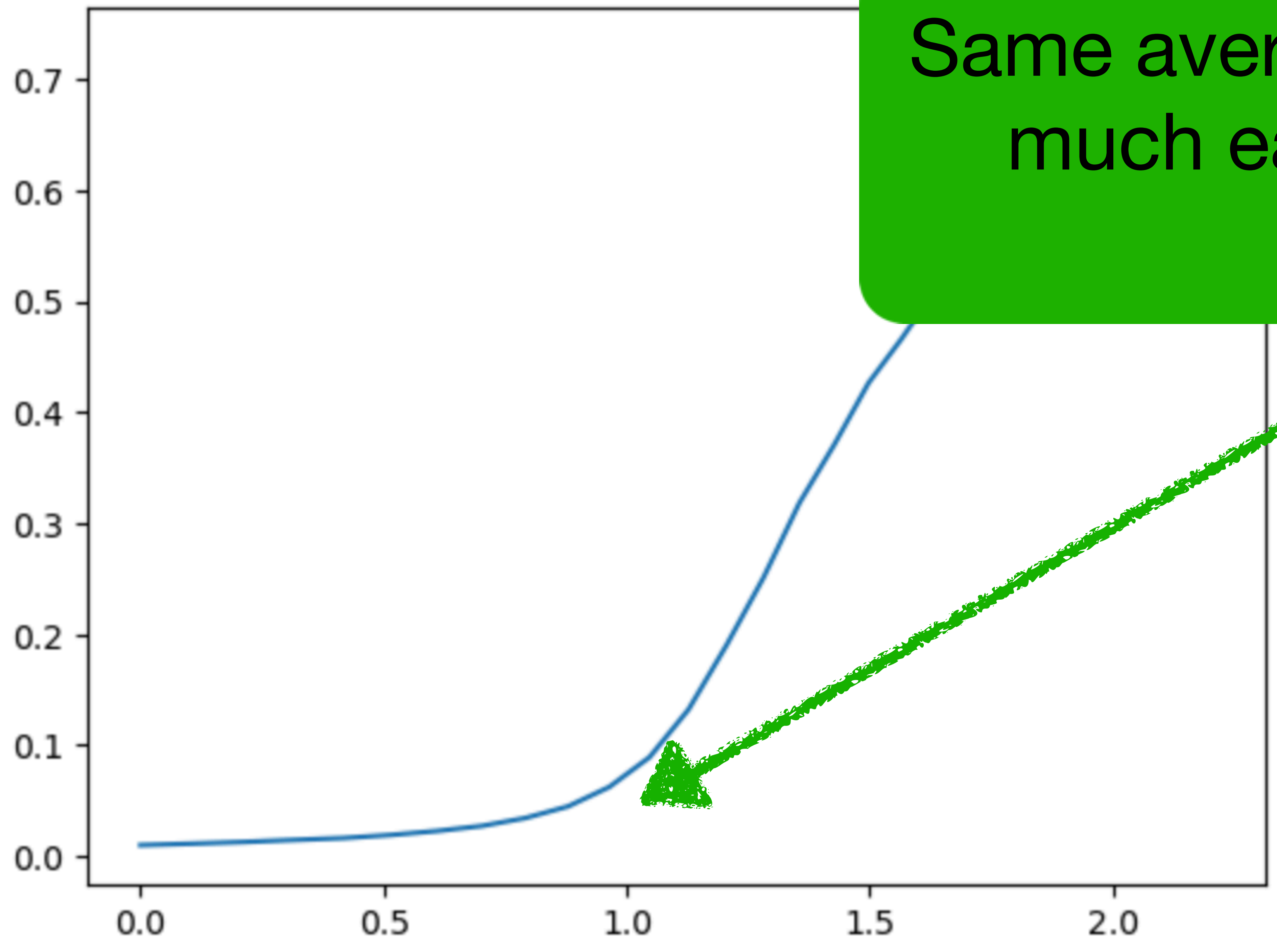
Barabási-Albert network



$R_0 = f\langle k \rangle$

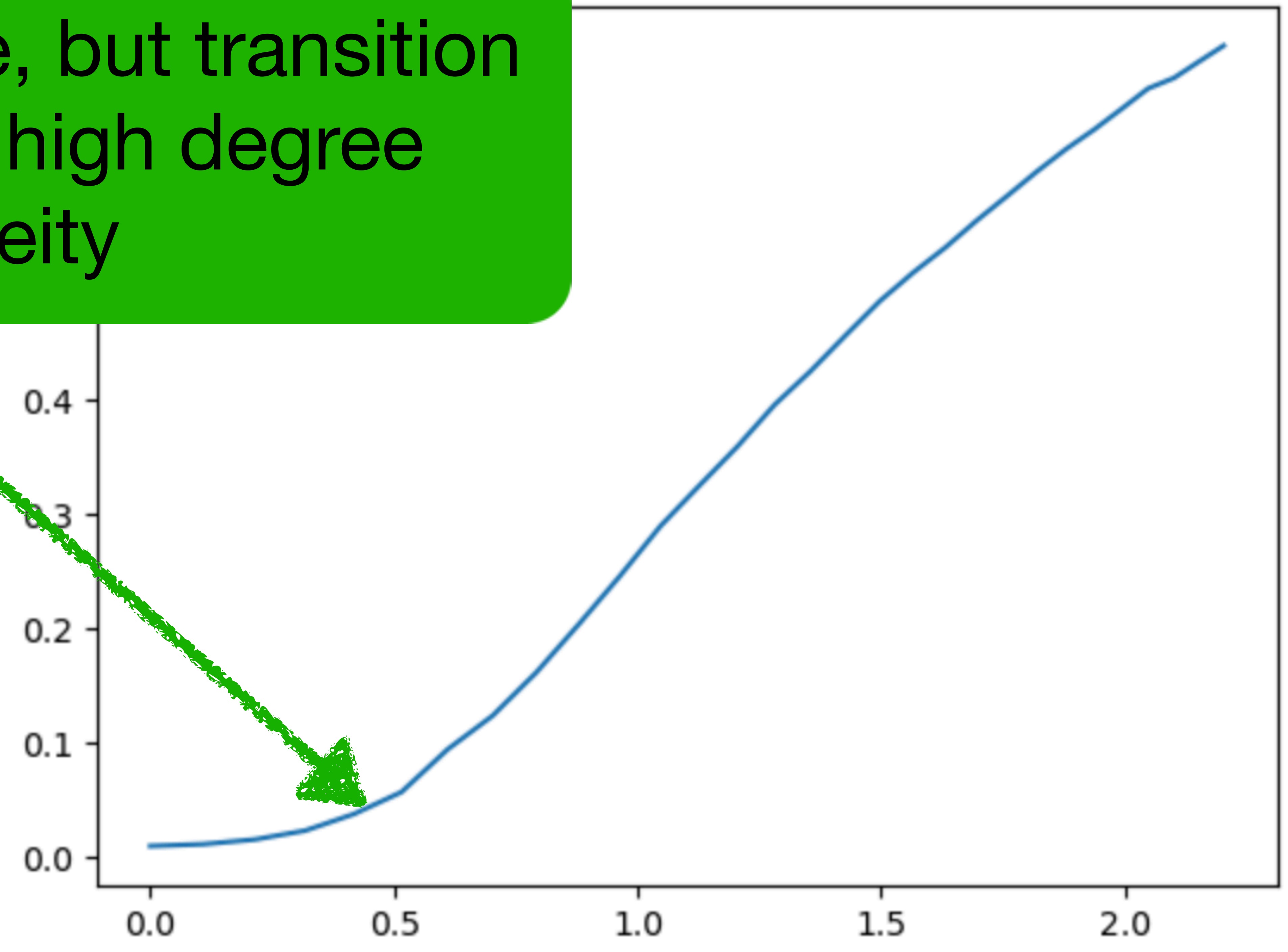
SIR spreading and phase transition

Fully random (ER) network



$R_0 = f\langle k \rangle$

Barabási-Albert network



$R_0 = f\langle k \rangle$

Same average degree, but transition much earlier when high degree heterogeneity

Opinion/behaviour spreading

$\neq$

infection/information spreading

Complex contagion

- “Simple contagion” of disease models is a good for spread of information

- Spread of behaviour more complex:

  - People will adopt behaviour only if they see enough of others adopting, especially if there are risks involved

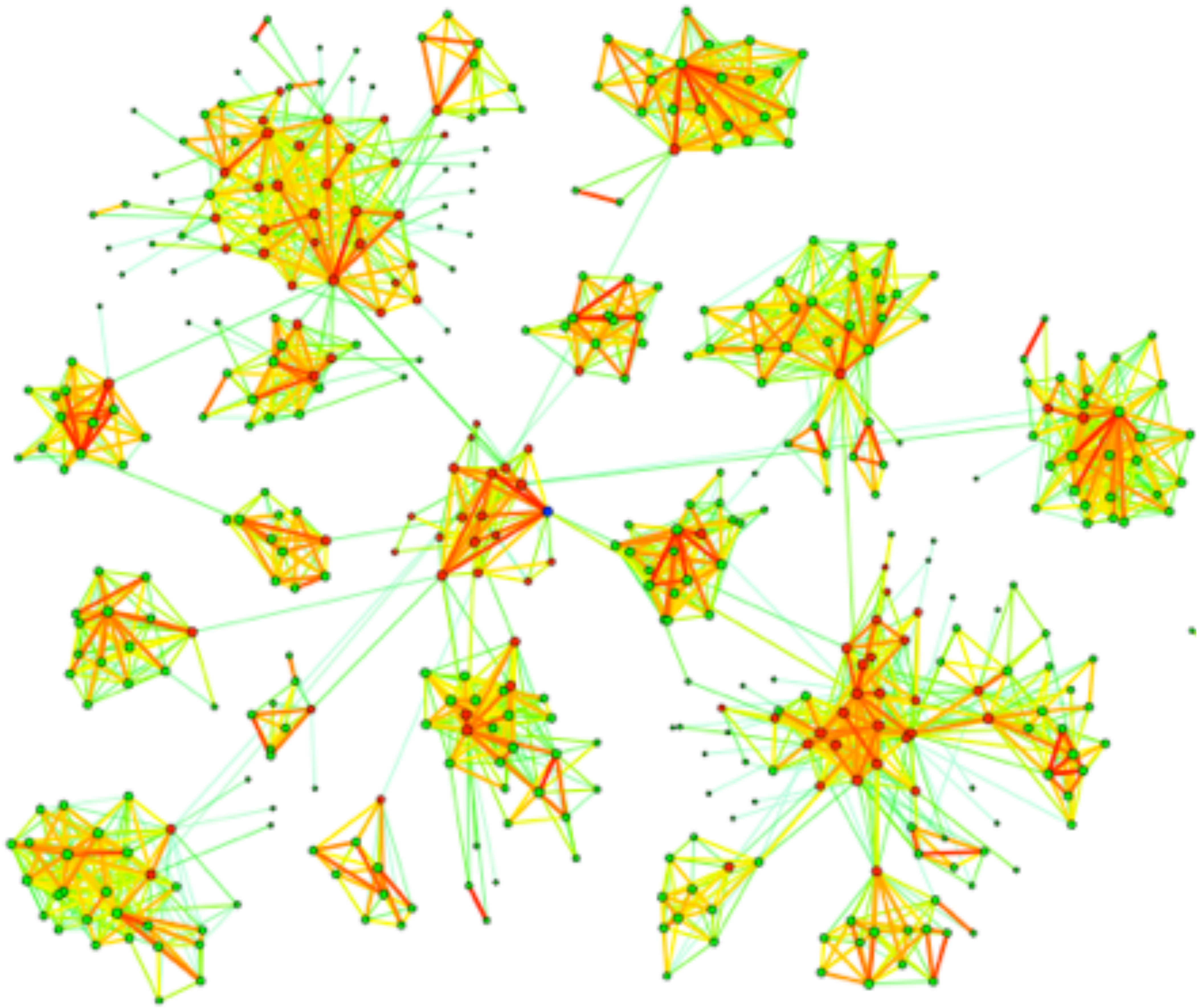
  - Examples: social movements, new fashion, new technologies

→ **Complex contagion**: e.g., a node needs exactly k neighbours to be adopters to adopt the behaviour themselves

Complex contagion: consequences

- Groups and clusters are important: adoption can spread within the groups

- Strength of weak ties (see lecture 5) is lost: individual links to far places in the network are not helpful for the spreading



Centola D., Macy M., Complex contagions and the weakness of long ties. Am. J. Sociol. 113, 702 (2007)

Complex contagion: an experiment

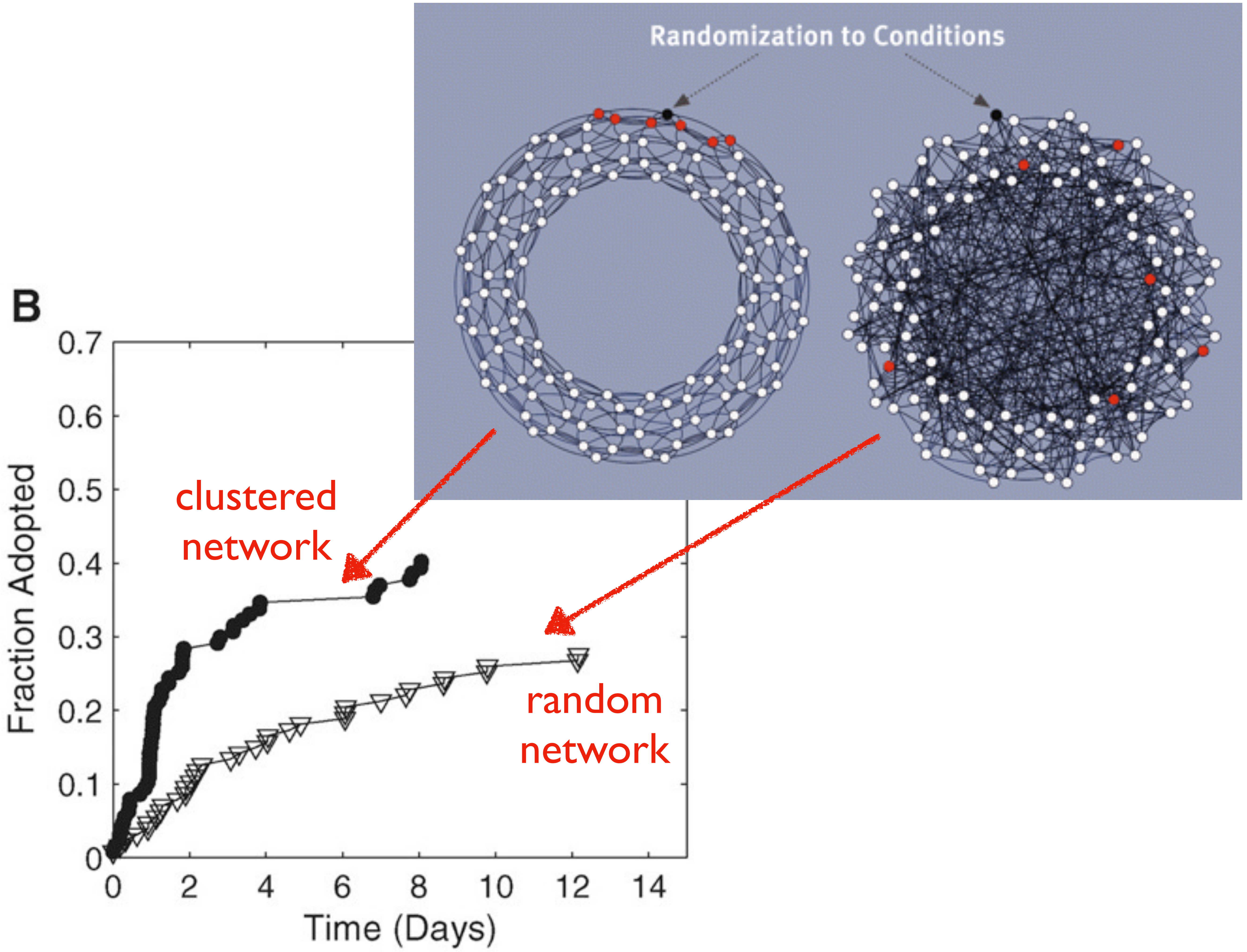
- Set-up: an online service, where users see what their “friends” do (but everyone is anonymous, just an user ID)

- The “friends” are determined by inserting users to a preset network (random or clustered); friends = neighbours in that network

- Some persons get a message that recommends registering to a health forum; if they do that, their “friends” see that they did it and may also imitate their behaviour

- “Individual adoption was much more likely when participants received social reinforcement from multiple neighbors in the social network.”

  The behavior spread farther and faster across clustered-lattice networks than across corresponding random networks.”



Centola D, The Spread of Behavior in an Online Social Network Experiment, Science 329, 1194 (2010)

Summary: dynamics on networks

- Network structure can a deciding effect on how dynamics unfold
- Social networks:
  - Skewed degree distributions: hubs are extremely important in any simple spreading processes
  - Clustering/group structure:
    - Hinder simple spreading and links between groups (weak ties) important
    - Can have opposite effects for complex contagion, links within groups (strong links) important

Network science courses at Aalto

Do you want to understand network theory at a deeper level?

→ CS-E5740 Complex Networks (periods I-II)

→ CS-E5745 Mathematical Methods in Network Science (period III)

→ CS-E5700 Hands-on Network Analysis (period IV)